CHAPTER 3

SPECIAL DEVICES

LEARNING OBJECTIVES

Upon completion of this chapter, you will be able to:

1. Explain the basic operation and the major applications of the Zener diode.
2. Describe the basic operation of the tunnel diode and the varactor.
3. Explain the basic operation of the silicon controlled rectifier and the TRIAC, and compare the advantages and disadvantages of each.
4. List the five most commonly used optoelectronic devices and explain the uses of each.
5. Describe the basic operation, applications, and major advantages of the unijunction transistor.
6. Describe the basic operation, applications, and major advantages of the field effect transistor and the metal oxide semiconductor field effect transistor.
7. Explain the basic operation and the major applications of the Zener diode.
8. Describe. the basic operation of the tunnel diode and the varactor.
9. Explain the basic operation of the silicon controlled rectifier and the TRIAC, and compare the advantages and disadvantages of each.
10. List the five most commonly used optoelectronic devices and explain the uses of each.
11. Describe the basic operation, applications, and major advantages of the unijunction transistor.
12. Describe the basic operation, applications, and major advantages of the field-effect transistor and the metal-oxide semiconductor field-effect transistor.

INTRODUCTION TO SPECIAL DEVICES

If you consider the sensitive nature and the various interacting properties of semiconductors, it should not be surprising to you that solid state devices can be designed for many different purposes. In fact, devices with special features are so numerous and new designs are so frequently introduced that it would be beyond the scope of this chapter to describe all of the devices in use today. Therefore, this chapter will include a variety of representative devices that are used extensively in Navy equipment to give you an idea of the diversity and versatility that have been made possible. These devices have been grouped into three categories: diodes, optoelectronic devices, and transistors. In this chapter each device will be described and the basic operation of each one will be discussed.

DIODES

Diodes are two terminal semiconductors of various types that are used in seemingly endless applications. The operation of normal PN-junction diodes has already been discussed, but there are a number of diodes with special properties with which you should be familiar. A discussion of all of the developments in the diode field would be impossible so some of the more commonly used special diodes have been selected for explanation. These include Zener diodes, tunnel diodes, varactors, silicon controlled rectifiers (SCR), and TRIACs.

Zener Diodes

When a PN-junction diode is reverse biased, the majority carriers (holes in the P-material and electrons in the N-material) move away from the junction. The barrier or depletion region becomes wider, as illustrated in figure 3-1, (view A, view B, view C) and majority carrier current flow becomes very difficult across the high resistance of the wide depletion region. The presence of minority carriers causes a small leakage current that remains nearly constant for all reverse voltages up to a certain value. Once this value has been exceeded, there is a sudden increase in the reverse current. The voltage at which the sudden increase in current occurs is called the BREAKDOWN VOLTAGE. At breakdown, the reverse current increases very rapidly with a slight increase in the reverse voltage. Any diode can be reverse biased to the point of breakdown, but not every diode can safely dissipate the power associated with breakdown. A Zener diode is a PN junction designed to operate in the reverse-bias breakdown region.

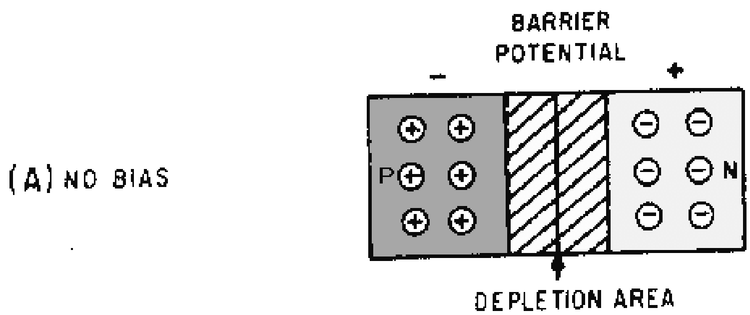


Figure 3-1A.—Effects of bias on the depletion region of a PN junction.

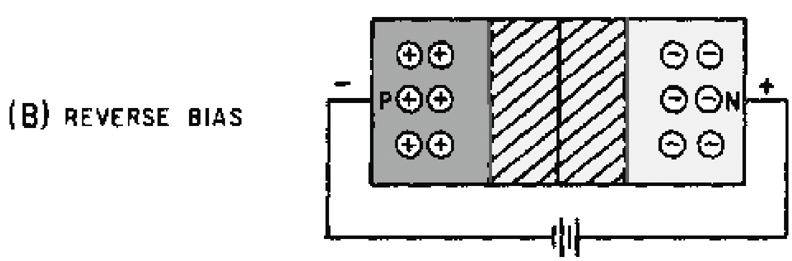


Figure 3-1B.—Effects of bias on the depletion region of a PN junction.

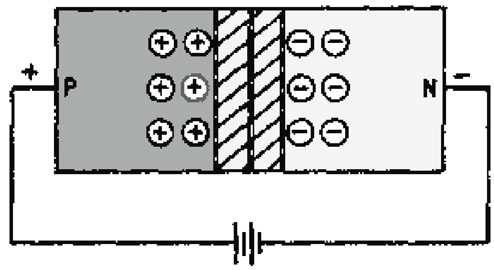


Figure 3-1C.—Effects of bias on the depletion region of a PN junction.

There are two distinct theories used to explain the behavior of PN junctions during breakdown: one is the ZENER EFFECT and the other is the AVALANCHE EFFECT.

The ZENER EFFECT was first proposed by Dr. Carl Zener in 1934. According to Dr. Zener's theory, electrical breakdown in solid dielectrics occurs by a process called QUANTUM-MECHANICAL TUNNELING. The Zener effect accounts for the breakdown below 5 volts; whereas, above 5 volts the breakdown is caused by the avalanche effect. Although the avalanche effect is now accepted as an explanation of diode breakdown, the term *Zener diode* is used to cover both types.

The true Zener effect in semiconductors can be described in terms of energy bands; however, only the two upper energy bands are of interest. The two upper bands, illustrated in figure 3-2, view A, are called the conduction band and the valence band.

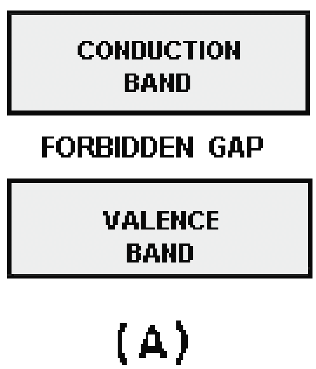


Figure 3-2A.-Energy diagram for Zener diode.

The CONDUCTION BAND is a band in which the energy level of the electrons is high enough that the electrons will move easily under the influence of an external field. Since current flow is the movement of electrons, the readily mobile electrons in the conduction band are capable of maintaining a current flow when an external field in the form of a voltage is applied. Therefore, solid materials that have many electrons in the conduction band are called conductors.

The VALENCE BAND is a band in which the energy level is the same as the valence electrons of the atoms. Since the electrons in these levels are attached to the atoms, the electrons are not free to move around as are the conduction band electrons. With the proper amount of energy added, however, the electrons in the valence band may be elevated to the conduction band energy level. To do this, the electrons must cross a gap that exists between the valence band energy level and the conduction band energy level. This gap is known as the FORBIDDEN ENERGY BAND or FORBIDDEN GAP. The energy difference across this gap determines whether a solid material will act as a conductor, a semiconductor, or an insulator.

A conductor is a material in which the forbidden gap is so narrow that it can be considered nonexistent. A semiconductor is a solid that contains a forbidden gap, as shown in figure 3-2, view A. Normally, a semiconductor has no electrons at the conduction band energy level. The energy provided by room temperature heat, however, is enough energy to overcome the binding force of a few valence electrons and to elevate them to the conduction band energy level. The addition of impurities to the semiconductor material increases both the number of free electrons in the conduction band and the number of electrons in the valence band that can be elevated to the conduction band. Insulators are materials in which the forbidden gap is so large that practically no electrons can be given enough energy to cross the gap. Therefore, unless extremely large amounts of heat energy are available, these materials will not conduct electricity.

View B of figure 3-2 is an energy diagram of a reverse-biased Zener diode. The energy bands of the P and N materials are naturally at different levels, but reverse bias causes the valence band of the P material to overlap the energy level of the conduction band in the N material. Under this condition, the valence electrons of the P material can cross the extremely thin junction region at the overlap point without acquiring any additional energy. This action is called tunneling. When the breakdown point of the PN junction is reached, large numbers of minority carriers "tunnel" across the junction to form the current that occurs at breakdown. The tunneling phenomenon only takes place in heavily doped diodes such as Zener diodes.

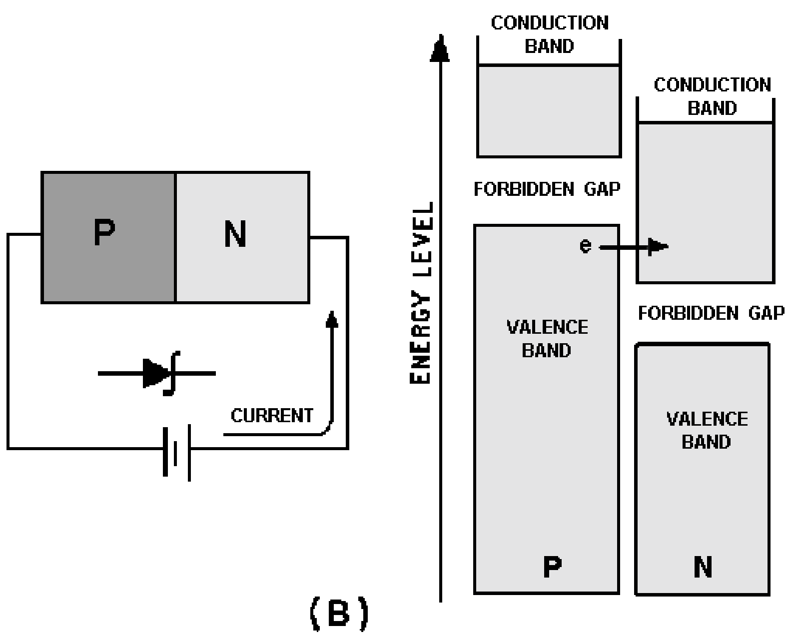


Figure 3-2B.-Energy diagram for Zener diode.

The second theory of reverse breakdown effect in diodes is known as AVALANCHE breakdown and occurs at reverse voltages beyond 5 volts. This type of breakdown diode has a depletion region that is deliberately made narrower than the depletion region in the normal PN-junction diode, but thicker than that in the Zener-effect diode. The thicker depletion region is achieved by decreasing the doping level from the level used in Zener-effect diodes. The breakdown is at a higher voltage because of the higher resistivity of the material. Controlling the doping level of the material during the manufacturing process can produce breakdown voltages ranging between about 2 and 200 volts.

The mechanism of avalanche breakdown is different from that of the Zener effect. In the depletion region of a PN junction, thermal energy is responsible for the formation of electron-hole pairs. The leakage current is caused by the movement of minority electrons, which is accelerated in the electric field across the barrier region. As the reverse voltage across the depletion region is increased, the reverse voltage eventually reaches a critical value. Once the critical or breakdown voltage has been reached, sufficient energy is gained by the thermally released minority electrons to enable the electrons to rupture covalent bonds as they collide with lattice atoms. The released electrons are also accelerated by the electric field, resulting in the release of further electrons, and so on, in a chain or avalanche effect. This process is illustrated in figure 3-3.

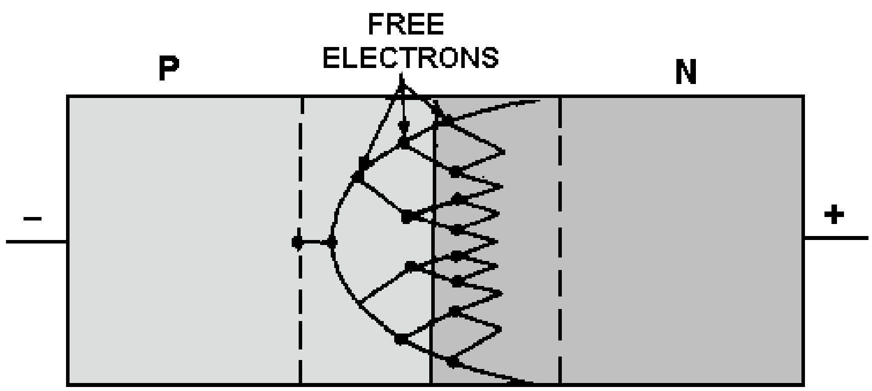


Figure 3-3.—Avalanche multiplication.

For reverse voltage slightly higher than breakdown, the avalanche effect releases an almost unlimited number of carriers so that the diode essentially becomes a short circuit. The current flow in this region is limited only by an external series current-limiting resistor. Operating a diode in the breakdown region does not damage it, as long as the maximum power dissipation rating of the diode is not exceeded. Removing the reverse voltage permits all carriers to return to their normal energy values and velocities.

Some of the symbols used to represent Zener diodes are illustrated in figure 3-4 (view A, view B, view C, view D, and view E). Note that the polarity markings indicate electron flow is with the arrow symbol instead of against it as in a normal PN-junction diode. This is because breakdown diodes are operated in the reverse-bias mode, which means the current flow is by minority current carriers.

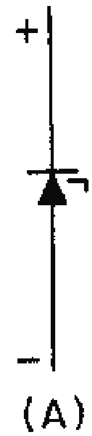


Figure 3-4A.—Schematic symbols for Zener diodes.

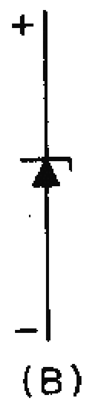


Figure 3-4B.—Schematic symbols for Zener diodes.

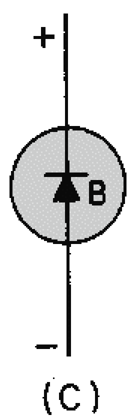


Figure 3-4C.—Schematic symbols for Zener diodes.

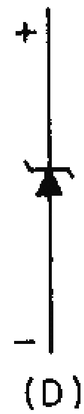


Figure 3-4D.—Schematic symbols for Zener diodes.

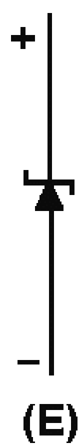


Figure 3-4E.—Schematic symbols for Zener diodes.

Zener diodes of various sorts are used for many purposes, but their most widespread use is as voltage regulators. Once the breakdown voltage of a Zener diode is reached, the voltage across the diode remains almost constant regardless of the supply voltage. Therefore they hold the voltage across the load at a constant level. This characteristic makes Zener diodes ideal voltage regulators, and they are found in almost all solid-state circuits in this capacity.

Q1. In a reverse biased PN-junction, which current carriers cause leakage current?

Q2. The action of a PN-junction during breakdown can be explained by what two theories?

Q3. Which breakdown theory explains the action that takes place in a heavily doped PN-junction with a reverse bias of less than 5 volts?

Q4. What is the doping level of an avalanche effect diode when compared to the doping level of a Zener-effect diode?

Q5. During avalanche effect breakdown, what limits current flow through the diode?

Q6. Why is electron flow with the arrow in the symbol of a Zener diode instead of against the arrow as it is in a normal diode?

The Tunnel Diode

In 1958, Leo Esaki, a Japanese scientist, discovered that if a semiconductor junction diode is heavily doped with impurities, it will have a region of negative resistance. The normal junction diode uses semiconductor materials that are lightly doped with one impurity atom for ten-million semiconductor atoms. This low doping level results in a relatively wide depletion region. Conduction occurs in the normal junction diode only if the voltage applied to it is large enough to overcome the potential barrier of the junction.

In the TUNNEL DIODE, the semiconductor materials used in forming a junction are doped to the extent of one-thousand impurity atoms for ten-million semiconductor atoms. This heavy doping produces an extremely narrow depletion zone similar to that in the Zener diode. Also because of the heavy doping, a tunnel diode exhibits an unusual current-voltage characteristic curve as compared with that of an ordinary junction diode. The characteristic curve for a tunnel diode is illustrated in figure 3-5.

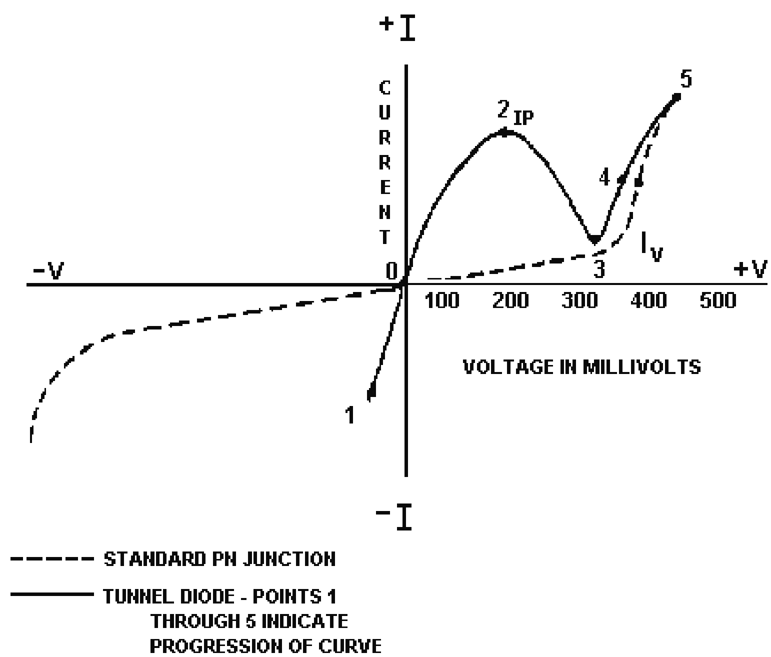


Figure 3-5.—Characteristic curve of a tunnel diode compared to that of a standard PN junction.

The three most important aspects of this characteristic curve are (1) the forward current increase to a peak (I_P) with a small applied forward bias, (2) the decreasing forward current with an increasing forward bias to a minimum valley current (I_V), and (3) the normal increasing forward current with further increases in the bias voltage. The portion of the characteristic curve between I_P and I_V is the region of negative resistance. An explanation of why a tunnel diode has a region of negative resistance is best understood by using energy levels as in the previous explanation of the Zener effect.

Simply stated the theory known as quantum-mechanical tunneling is an electron crossing a PN-junction without having sufficient energy to do so otherwise. Because of the heavy doping the width of the depletion region is only one-millionth of an inch. You might think of the process simply as an arc-over between the N- and the P-side across the depletion region.

Figure 3-6 shows the equilibrium energy level diagram of a tunnel diode with no bias applied. Note in view A that the valence band of the P-material overlaps the conduction band of the N-material. The majority electrons and holes are at the same energy level in the equilibrium state. If there is any movement of current carriers across the depletion region due to thermal energy, the net current flow will be zero because equal numbers of current carriers flow in opposite directions. The zero net current flow is marked by a "0" on the current-voltage curve illustrated in view B.

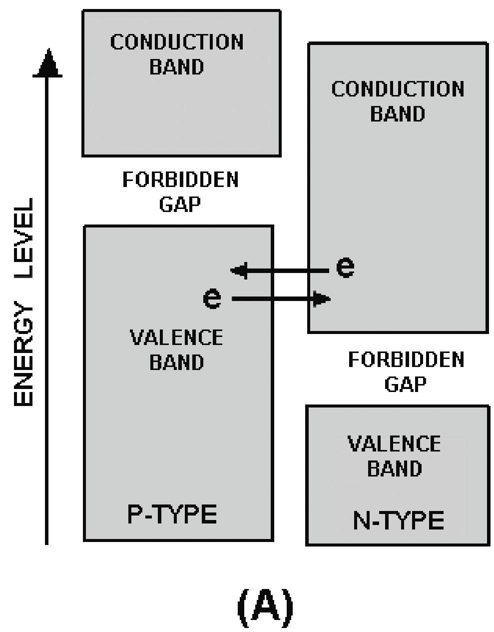


Figure 3-6A.—Tunnel diode energy diagram with no bias.

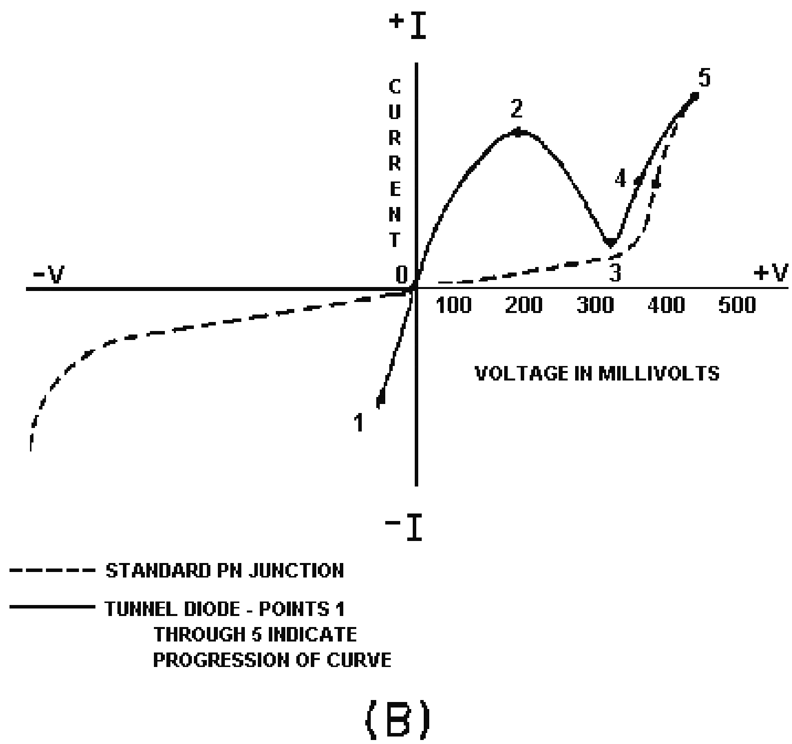


Figure 3-6B.—Tunnel diode energy diagram with no bias.

Figure 3-7, view A, shows the energy diagram of a tunnel diode with a small forward bias (50 millivolts) applied. The bias causes unequal energy levels between some of the majority carriers at the energy band overlap point, but not enough of a potential difference to cause the carriers to cross the forbidden gap in the normal manner. Since the valence band of the P-material and the conduction band of the N-material still overlap, current carriers tunnel across at the overlap and cause a substantial current flow. The amount of current flow is marked by point 2 on the curve in view B. Note in view A that the amount of overlap between the valence band and the conduction band decreased when forward bias was applied.

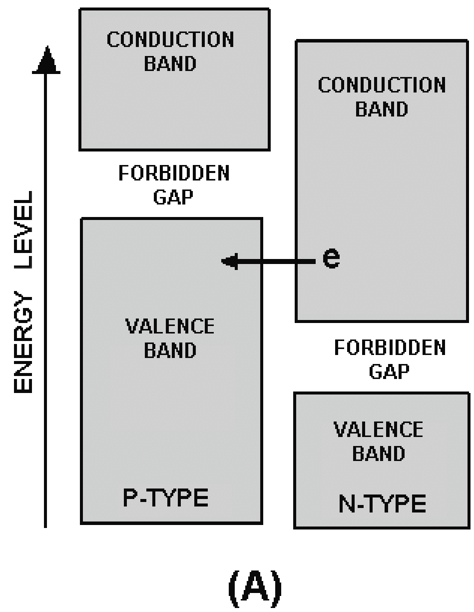


Figure 3-7A.—Tunnel diode energy diagram with 50 millivolts bias.

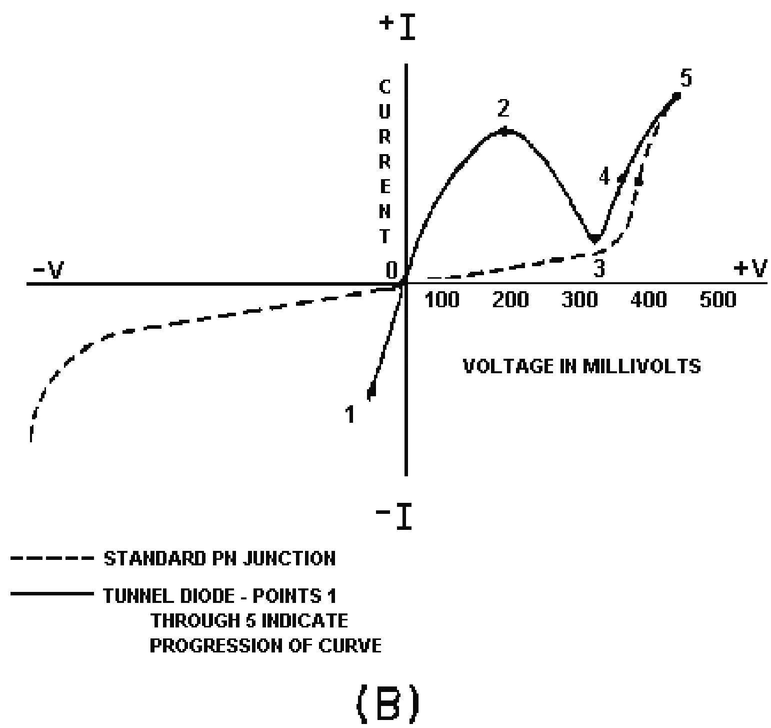


Figure 3-7B.—Tunnel diode energy diagram with 50 millivolts bias.

Figure 3-8, view A, is the energy diagram of a tunnel diode in which the forward bias has been increased to 450 millivolts. As you can see, the valence band and the conduction band no longer overlap at this point, and tunneling can no longer occur. The portion of the curve in view B from point 2 to point 3 shows the decreasing current that occurs as the bias is increased, and the area of overlap becomes smaller. As the overlap between the two energy bands becomes smaller, fewer and fewer electrons can tunnel across the junction. The portion of the curve between point 2 and point 3 in which current decreases as the voltage increases is the negative resistance region of the tunnel diode.

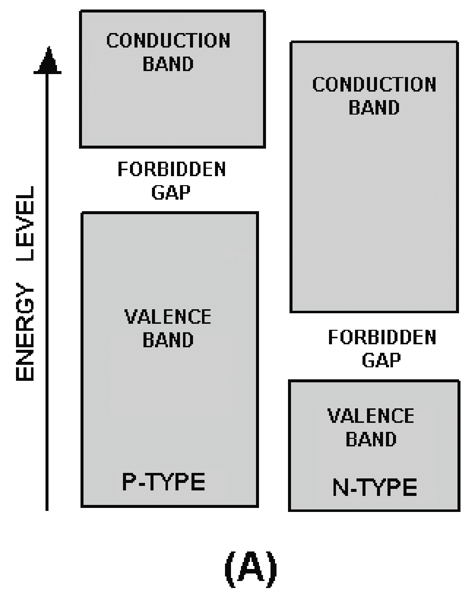


Figure 3-8A.—Tunnel diode energy diagram with 450 millivolts bias.

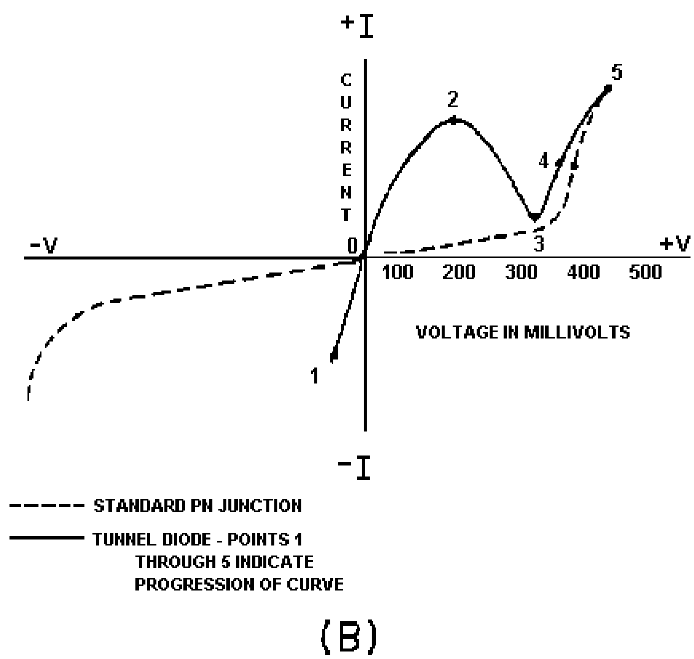


Figure 3-8B.—Tunnel diode energy diagram with 450 millivolts bias.

Figure 3-9, view A, is the energy diagram of a tunnel diode in which the forward bias has been increased even further. The energy bands no longer overlap and the diode operates in the same manner as a normal PN junction, as shown by the portion of the curve in view (B) from point 3 to point 4.

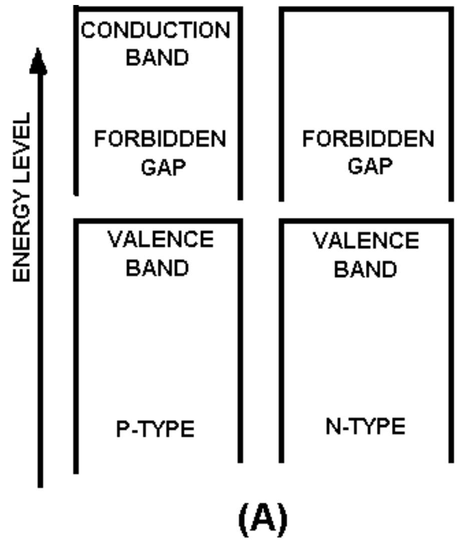


Figure 3-9A.—Tunnel diode energy diagram with 600 millivolts bias.

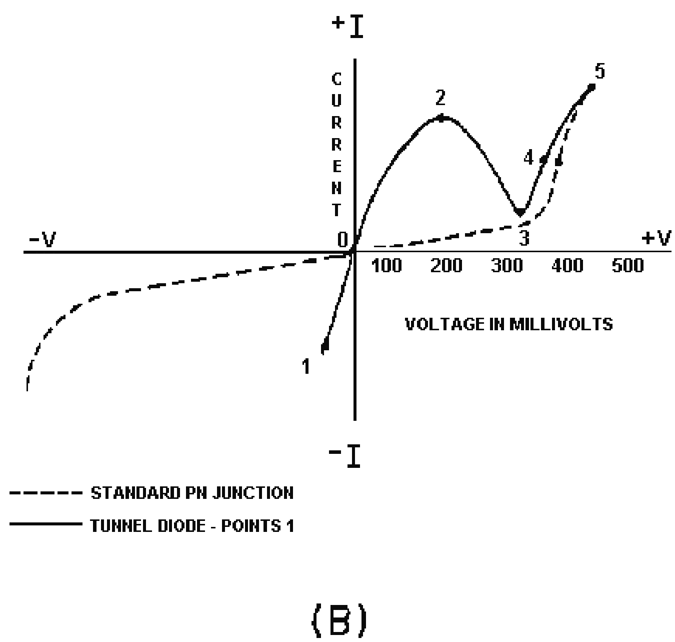


Figure 3-9B.—Tunnel diode energy diagram with 600 millivolts bias.

The negative resistance region is the most important and most widely used characteristic of the tunnel diode. A tunnel diode biased to operate in the negative resistance region can be used as either an oscillator or an amplifier in a wide range of frequencies and applications. Very high frequency applications using the tunnel diode are possible because the tunneling action occurs so rapidly that there is no transit time effect and therefore no signal distortion. Tunnel diodes are also used extensively in high-speed switching circuits because of the speed of the tunneling action.

Several schematic symbols are used to indicate a tunnel diode. These symbols are illustrated in figure 3-10 (view A, view B, view C, and view D).

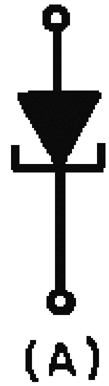


Figure 3-10A.—Tunnel diode schematic symbols.

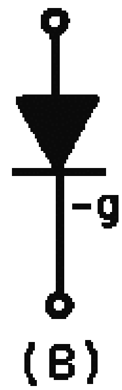


Figure 3-10B.—Tunnel diode schematic symbols.

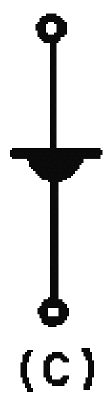


Figure 3-10C.—Tunnel diode schematic symbols.

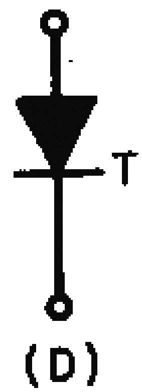


Figure 3-10D.—Tunnel diode schematic symbols.

Varactor

The VARACTOR, or varicap, as the schematic drawing in figure 3-11 suggests, is a diode that behaves like a variable capacitor, with the PN junction functioning like the dielectric and plates of a common capacitor. Understanding how the varactor operates is an important prerequisite to understanding field-effect transistors, which will be covered later in this topic.

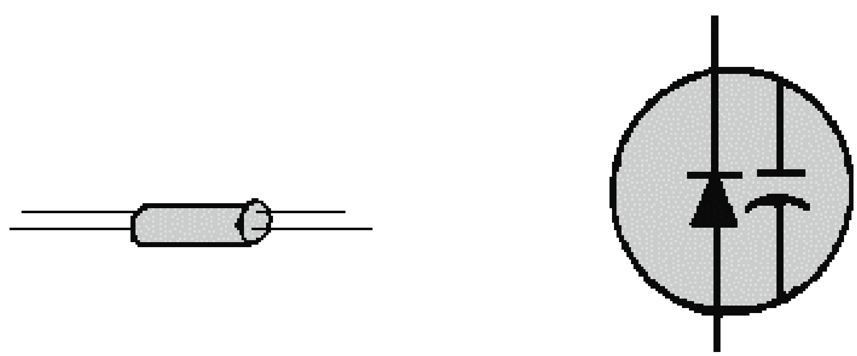


Figure 3-11.—Varactor diode.

Figure 3-12 shows a PN junction. Surrounding the junction of the P and N materials is a narrow region void of both positively and negatively charged current carriers. This area is called the depletion region.

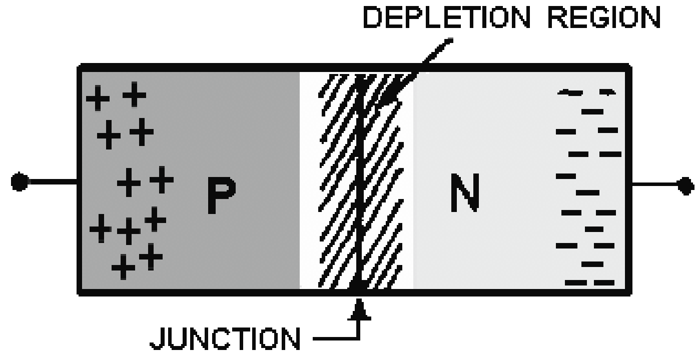


Figure 3-12.—PN junction.

The size of the depletion region in a varactor diode is directly related to the bias. Forward biasing makes the region smaller by repelling the current carriers toward the PN junction. If the applied voltage is large enough (about .5 volt for silicon material), the negative particles will cross the junction and join with the positive particles, as shown in figure 3-13. This forward biasing causes the depletion region to decrease, producing a low resistance at the PN junction and a large current flow across it. This is the condition for a forward-biased diode. On the other hand, if reverse-bias voltage is applied to the PN junction, the size of its depletion region increases as the charged particles on both sides move away from the junction. This condition, shown in figure 3-14, produces a high resistance between the terminals and allows little current flow (only in the microampere range). This is the operating condition for the varactor diode, which is nothing more than a special PN junction.

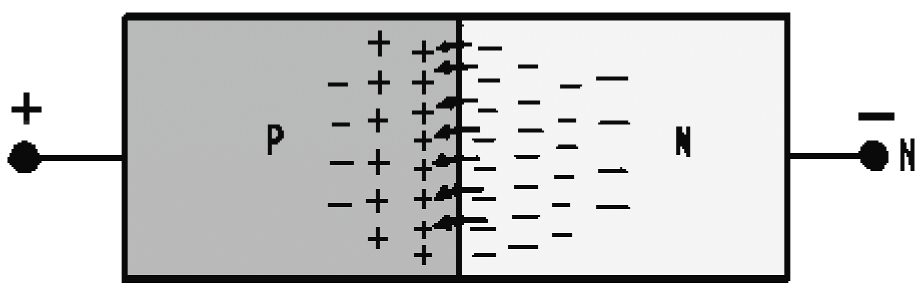


Figure 3-13.—Forward-biased PN junction.

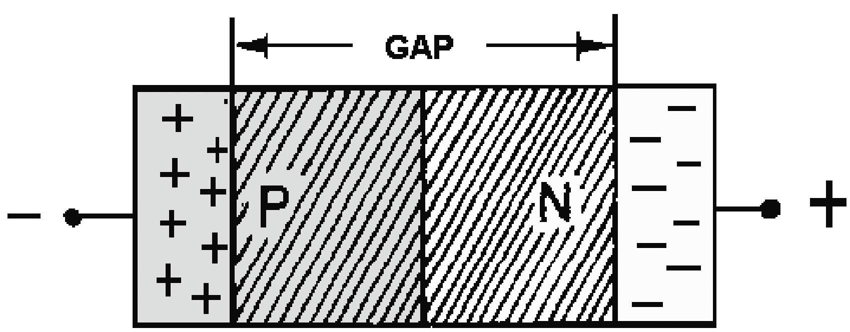


Figure 3-14.—Reverse-biased PN junction.

As the figure shows, the insulation gap formed by reverse biasing of the varactor is comparable to the layer of dielectric material between the plates of a common capacitor. Furthermore, the formula used to calculate capacitance

$$C = \frac{AK}{d}$$

Where

A = plate area

K = a constant value

d = distance between plates

can be applied to both the varactor and the capacitor. In this case, the size of the insulation gap of the varactor, or depletion region, is substituted for the distance between the plates of the capacitor. By varying the reverse-bias voltage applied to the varactor, the width of the "gap" may be varied. An increase in reverse bias increases the width of the gap (d) which reduces the capacitance (C) of the PN junction. Therefore, the capacitance of the varactor is inversely proportional to the applied reverse bias.

The ratio of varactor capacitance to reverse-bias voltage change may be as high as 10 to 1. Figure 3-15 shows one example of the voltage-to-capacitance ratio. View A shows that a reverse bias of 3 volts produces a capacitance of 20 picofarads in the varactor. If the reverse bias is increased to 6 volts, as shown in view B, the depletion region widens and capacitance drops to 5 picofarads. Each 1-volt increase in bias voltage causes a 5-picofarad decrease in the capacitance of the varactor; the ratio of change is therefore 5 to 1. Of course any decrease in applied bias voltage would cause a proportionate increase in capacitance, as the depletion region narrows. Notice that the value of the capacitance is small in the picofarad range.

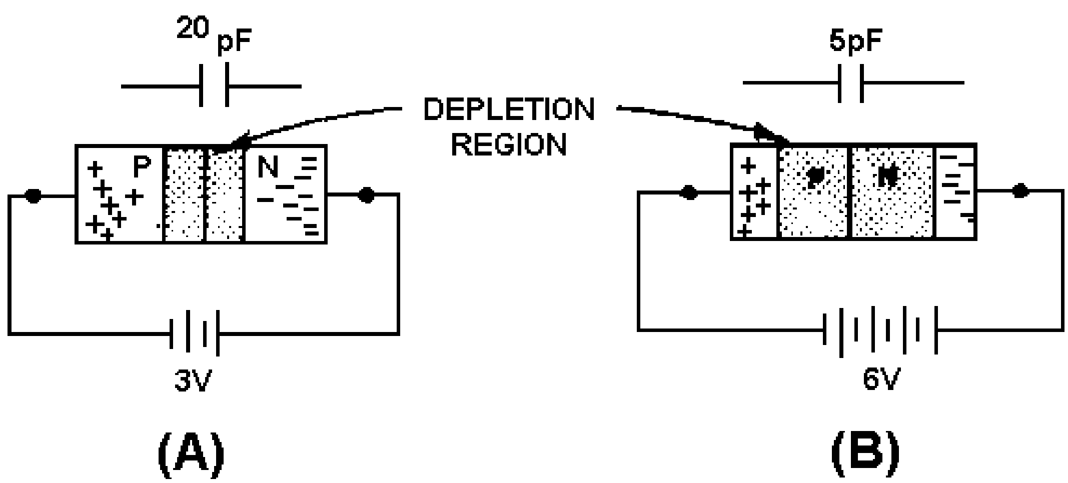


Figure 3-15.—Varactor capacitance versus bias voltage.

In general, varactors are used to replace the old style variable capacitor tuning. They are used in tuning circuits of more sophisticated communication equipment and in other circuits where variable capacitance is required. One advantage of the varactor is that it allows a dc voltage to be used to tune a circuit for simple remote control or automatic tuning functions. One such application of the varactor is as a variable tuning capacitor in a receiver or transmitter tank circuit like that shown in figure 3-16.

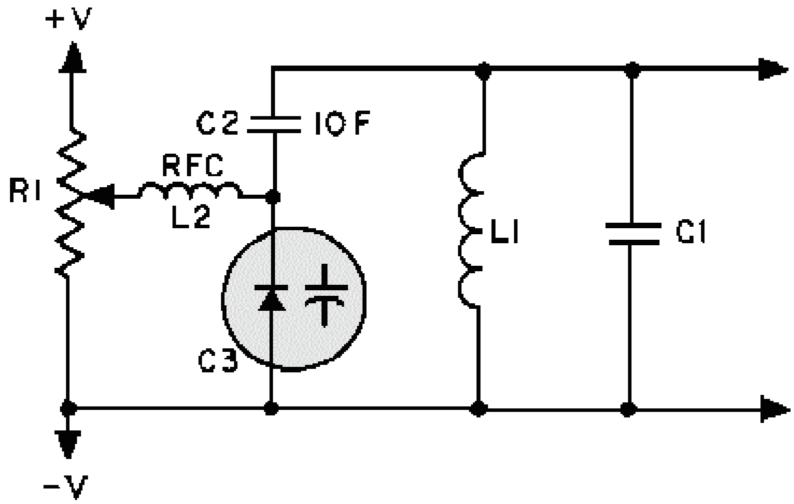


Figure 3-16.—Varactor tuned resonant circuit.

Figure 3-16 shows a dc voltage felt at the wiper of potentiometer R1 which can be adjusted between +V and –V. The dc voltage, passed through the low resistance of radio frequency choke L2, acts to reverse bias varactor diode C3. The capacitance of C3 is in series with C2, and the equivalent capacitance of C2 and C3 is in parallel with tank circuit L1-C1. Therefore, any variation in the dc voltage at R1 will vary both the capacitance of C3 and the resonant frequency of the tank circuit. The radio-frequency choke provides high inductive reactance at the tank frequency to prevent tank loading by R1. C2 acts to block dc from the tank as well as to fix the tuning range of C3.

An ohmmeter can be used to check a varactor diode in a circuit. A high reverse-bias resistance and a low forward-bias resistance with a 10 to 1 ratio in reverse-bias to forward-bias resistance is considered normal.

Q7. What is the main difference in construction between normal PN junction diodes and tunnel diodes?

Q8. What resistance property is found in tunnel diodes but not in normal diodes?

Q9. When compared to the ordinary diode, the tunnel diode has what type of depletion region?

Q10. In the tunnel diode, the tunneling current is at what level when the forbidden gap of the N-type material is at the same energy level as the empty states of the P-type material?

Q11. The varactor displays what useful electrical property?

Q12. When a PN junction is forward biased, what happens to the depletion region?

Q13. When the reverse bias on a varactor is increased, what happens to the effective capacitance?

Silicon Controlled Rectifier (SCR)

The SILICON CONTROLLED RECTIFIER, usually referred to as an SCR, is one of the family of semiconductors that includes transistors and diodes. A drawing of an SCR and its schematic representation is shown in views A and B of figure 3-17. Not all SCRs use the casing shown, but this is typical of most of the high-power units.

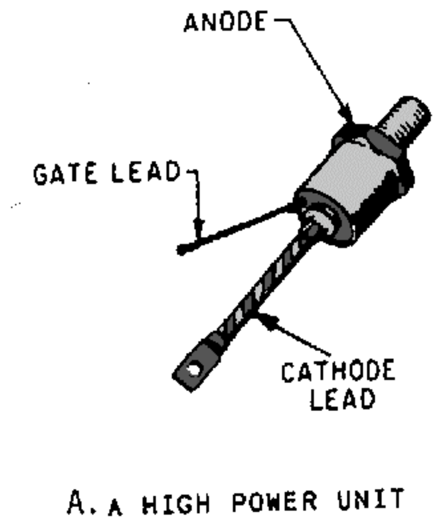


Figure 3-17A.—Silicon controlled rectifier.

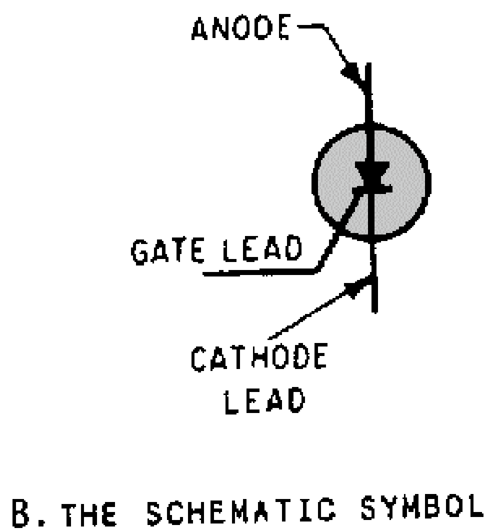


Figure 3-17B.—Silicon controlled rectifier.

Although it is not the same as either a diode or a transistor, the SCR combines features of both. Circuits using transistors or rectifier diodes may be greatly improved in some instances through the use of SCRs.

The basic purpose of the SCR is to function as a switch that can turn on or off small or large amounts of power. It performs this function with no moving parts that wear out and no points that require replacing. There can be a tremendous power gain in the SCR; in some units a very small triggering current is able to switch several hundred amperes without exceeding its rated abilities. The SCR can often replace much slower and larger mechanical switches. It even has many advantages over its more complex and larger electron tube equivalent, the thyratron.

The SCR is an extremely fast switch. It is difficult to cycle a mechanical switch several hundred times a minute; yet, some SCRs can be switched 25,000 times a second. It takes just microseconds (millionths of a second) to turn on or off these units. Varying the time that a switch is on as compared to the time that it is off regulates the amount of power flowing through the switch. Since most devices can operate on pulses of power (alternating current is a special form of alternating positive and negative pulse), the SCR can be used readily in control applications. Motor-speed controllers, inverters, remote switching units, controlled rectifiers, circuit overload protectors, latching relays, and computer logic circuits all use the SCR.

The SCR is made up of four layers of semiconductor material arranged PNPN. The construction is shown in view A of figure 3-18. In function, the SCR has much in common with a diode, but the theory of operation of the SCR is best explained in terms of transistors.

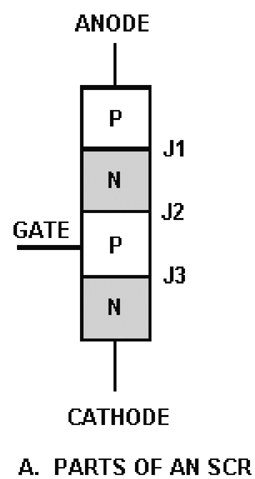


Figure 3-18A.—SCR structure.

Consider the SCR as a transistor pair, one PNP and the other NPN, connected as shown in views B and C. The anode is attached to the upper P-layer; the cathode, C, is part of the lower N-layer; and the gate terminal, G, goes to the P-layer of the NPN triode.

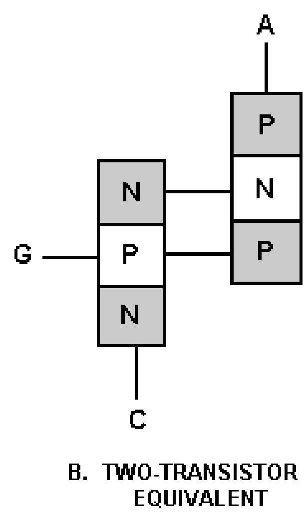
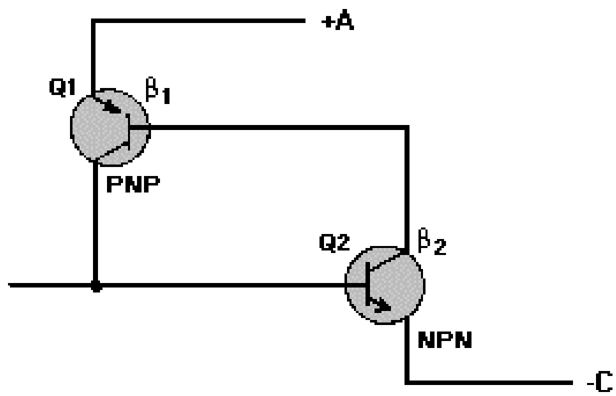


Figure 3-18B.—SCR structure.



C. TWO-TRANSISTOR SCHEMATIC

Figure 3-18C.—SCR structure.

In operation the collector of Q2 drives the base of Q1, while the collector of Q1 feeds back to the base of Q2. (Beta) 1 is the current gain of Q1, and (Beta) 2 is the current gain of Q2. The gain of this positive feedback loop is their product, 1 times 2. When the product is less than one, the circuit is stable; if the product is greater than unity, the circuit is regenerative. A small negative current applied to terminal G will bias the NPN transistor into cutoff, and the loop gain is less than unity. Under these conditions, the only current that can exist between output terminals A and C is the very small cutoff collector current of the two transistors. For this reason the impedance between A and C is very high.

When a positive current is applied to terminal G, transistor Q2 is biased into conduction, causing its collector current to rise. Since the current gain of Q2 increases with increased collector current, a point (called the breakover point) is reached where the loop gain equals unity and the circuit becomes regenerative. At this point, collector current of the two transistors rapidly increases to a value limited only by the external circuit. Both transistors are driven into saturation, and the impedance between A and C is very low. The positive current applied to terminal G, which served to trigger the self-regenerative action, is no longer required since the collector of PNP transistor Q1 now supplies more than enough current to drive Q2. The circuit will remain on until it is turned off by a reduction in the collector current to a value below that necessary to maintain conduction.

The characteristic curve for the SCR is shown in figure 3-19. With no gate current, the leakage current remains very small as the forward voltage from cathode to anode is increased until the breakdown point is reached. Here the center junction breaks down, the SCR begins to conduct heavily, and the drop across the SCR becomes very low.

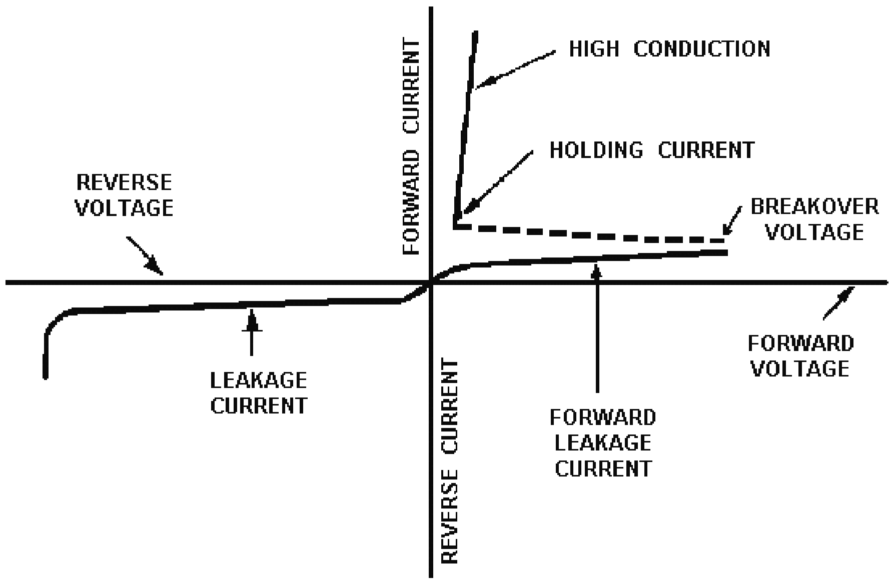


Figure 3-19.—Characteristic curve for an SCR.

The effect of a gate signal on the firing of an SCR is shown in figure 3-20. Breakdown of the center junction can be achieved at speeds approaching a microsecond by applying an appropriate signal to the gate lead, while holding the anode voltage constant. After breakdown, the voltage across the device is so low that the current through it from cathode to anode is essentially determined by the load it is feeding.

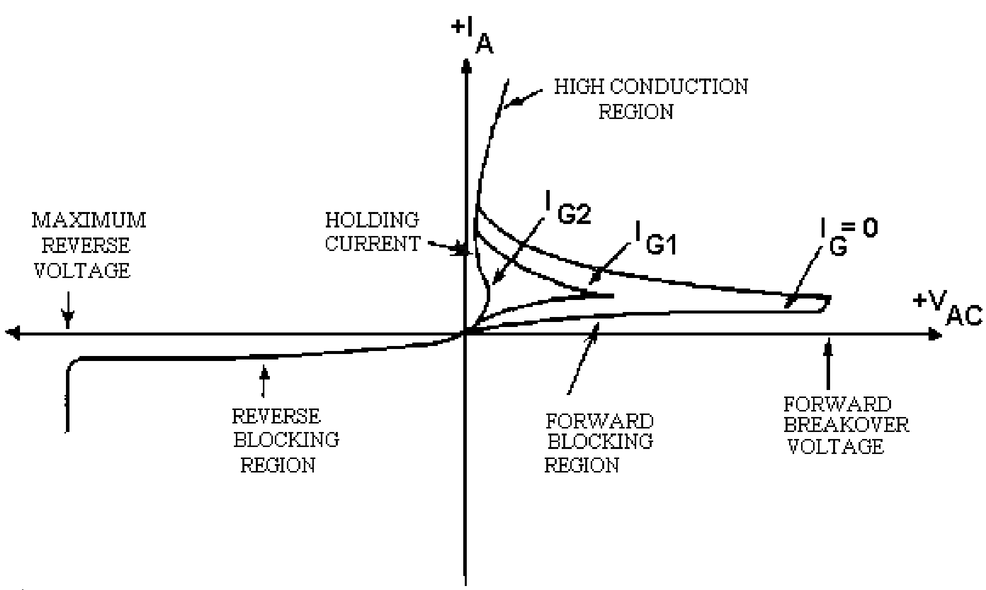


Figure 3-20.—SCR characteristic curve with various gate signals.

The important thing to remember is that a small current from gate to cathode can fire or trigger the SCR, changing it from practically an open circuit to a short circuit. The only way to change it back again (to commutate it) is to reduce the load current to a value less than the minimum forward-bias current. Gate current is required only until the anode current has completely built up to a point sufficient to sustain conduction (about 5 microseconds in resistive-load circuits). After conduction from cathode to anode begins, removing the gate current has no effect.

The basic operation of the SCR can be compared to that of the thyratron. The thyratron is an electron tube, normally gas filled, that uses a filament or a heater. The SCR and the thyratron function in a very similar manner. Figure 3-21 shows the schematic of each with the corresponding elements labeled. In both types of devices, control by the input signal is lost after they are triggered. The control grid (thyratron) and the gate (SCR) have no further effect on the magnitude of the load current after conduction begins. The load current can be interrupted by one or more of three methods: (1) the load circuit must be opened by a switch, (2) the plate (anode) voltage must be reduced below the ionizing potential of the gas (thyratron), (3) the forward-bias current must be reduced below a minimum value required to sustain conduction (SCR). The input resistance of the SCR is relatively low (approximately 100 ohms) and requires a current for triggering; the input resistance of the thyratron is exceptionally high, and requires a voltage input to the grid for triggering action.

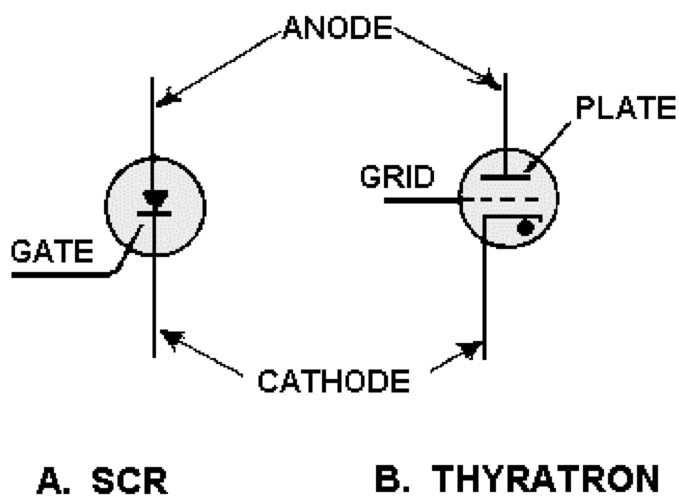


Figure 3-21.—Comparison of an SCR and a thyratron.

The applications of the SCR as a rectifier are many. In fact, its many applications as a rectifier give this semiconductor device its name. When alternating current is applied to a rectifier, only the positive or negative halves of the sine wave flow through. All of each positive or negative half cycle appears in the output. When an SCR is used, however, the controlled rectifier may be turned on at any time during the half cycle, thus controlling the amount of dc power available from zero to maximum, as shown in figure 3-22. Since the output is actually dc pulses, suitable filtering can be added if continuous direct current is needed. Thus any dc operated device can have controlled amounts of power applied to it. Notice that the SCR must be turned on at the desired time for each cycle.

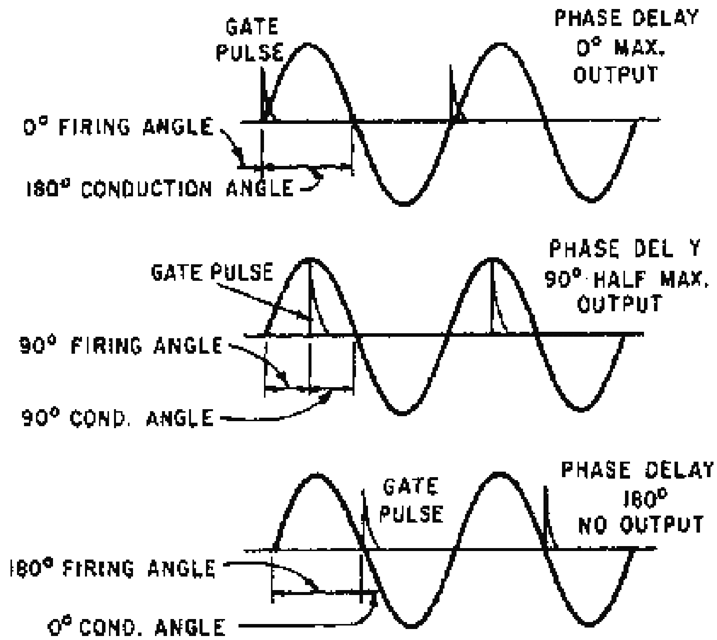


Figure 3-22.—SCR gate control signals.

When an ac power source is used, the SCR is turned off automatically, since current and voltage drop to zero every half cycle. By using one SCR on positive alternations and one on negative, full-wave rectification can be accomplished, and control is obtained over the entire sine wave. The SCR serves in this application just as its name implies—as a controlled rectifier of ac voltage.

Q14. The SCR is primarily used for what function?

Q15. When an SCR is forward biased, what is needed to cause it to conduct?

Q16. What is the only way to cause an SCR to stop conducting?

TRIAC

The TRIAC is a three-terminal device similar in construction and operation to the SCR. The TRIAC controls and conducts current flow during both alternations of an ac cycle, instead of only one. The schematic symbols for the SCR and the TRIAC are compared in figure 3-23. Both the SCR and the TRIAC have a gate lead. However, in the TRIAC the lead on the same side as the gate is "main terminal 1," and the lead opposite the gate is "main terminal 2." This method of lead labeling is necessary because the TRIAC is essentially two SCRs back to back, with a common gate and common terminals. Each terminal is, in effect, the anode of one SCR and the cathode of another, and either terminal can receive an input. In fact, the functions of a TRIAC can be duplicated by connecting two actual SCRs as shown in figure 3-24. The result is a three-terminal device identical to the TRIAC. The common anode-cathode connections form main terminals 1 and 2, and the common gate forms terminal 3.

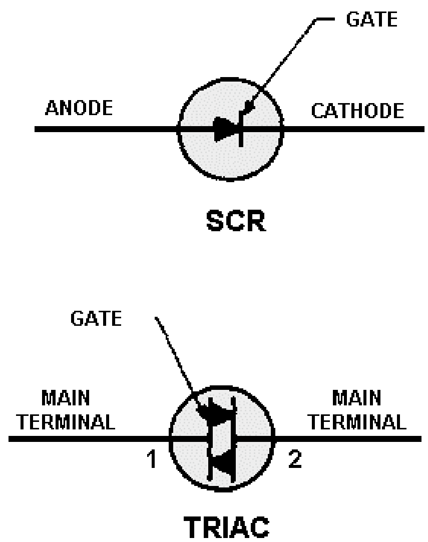


Figure 3-23.—Comparison of SCR and TRIAC symbols.

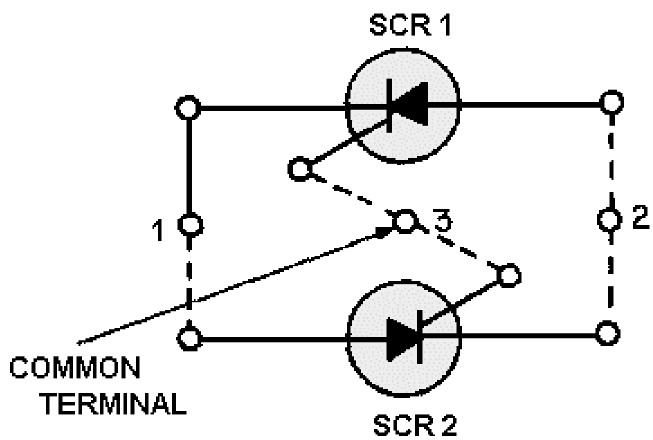


Figure 3-24.—Back to back SCR equivalent circuit.

The difference in current control between the SCR and the TRIAC can be seen by comparing their operation in the basic circuit shown in figure 3-25.

In the circuit shown in view A, the SCR is connected in the familiar half-wave arrangement. Current will flow through the load resistor (R_L) for one alternation of each input cycle. Diode CR1 is necessary to ensure a positive trigger voltage.

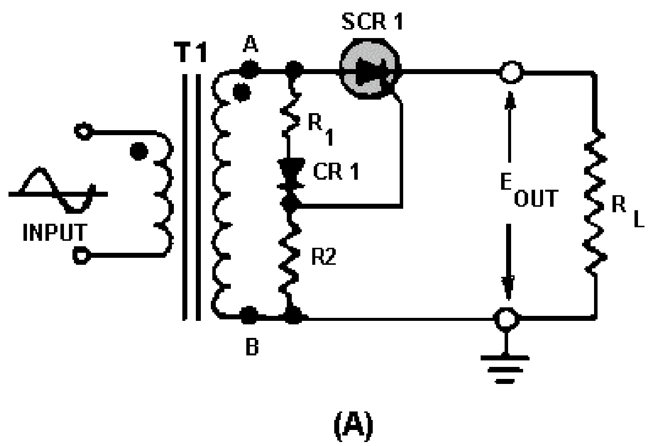


Figure 3-25A.—Comparison of SCR and TRIAC circuits.

In the circuit shown in view B, with the TRIAC inserted in the place of the SCR, current flows through the load resistor during both alternations of the input cycle. Because either alternation will trigger the gate of the TRIAC, CR1 is not required in the circuit. Current flowing through the load will reverse direction for half of each input cycle. To clarify this difference, a comparison of the waveforms seen at the input, gate, and output points of the two devices is shown in figure 3-26.

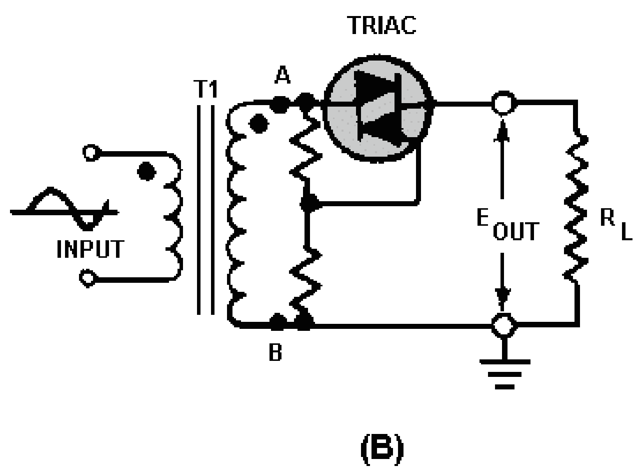


Figure 3-25B.—Comparison of SCR and TRIAC circuits.

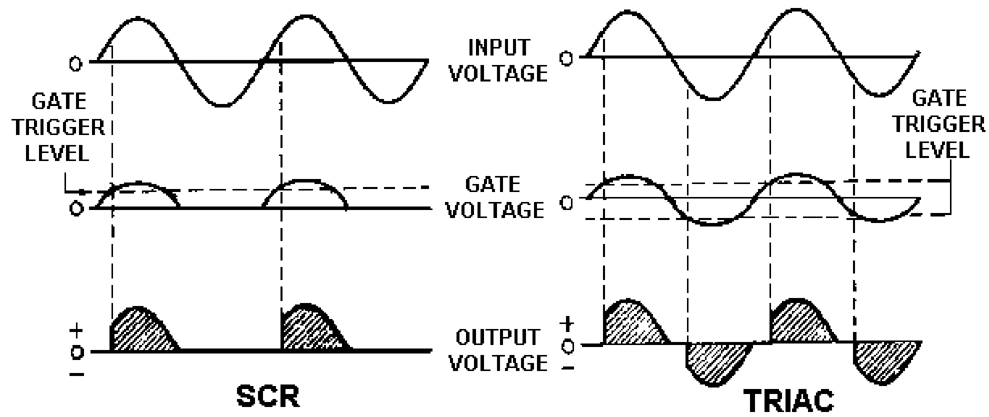


Figure 3-26.—Comparison of SCR and TRIAC waveforms.

Q17. The TRIAC is similar in operation to what device?

Q18. When used for ac current control, during which alternation of the ac cycle does the TRIAC control current flow?

Optoelectronic Devices

OPTOELECTRONIC devices either produce light or use light in their operation. The first of these, the light-emitting diode (LED), was developed to replace the fragile, short-life incandescent light bulbs used to indicate on/off conditions on panels. A LIGHT-EMITTING DIODE is a diode which, when forward biased, produces visible light. The light may be red, green, or amber, depending upon the material used to make the diode.

Figure 3-27 shows an LED and its schematic symbol. The LED is designated by a standard diode symbol with two arrows pointing away from the cathode. The arrows indicate light leaving the diode. The circuit symbols for all optoelectronic devices have arrows pointing either toward them, if they <u>use</u> light, or away from them, if they <u>produce</u> light. The LED operating voltage is small, about 1.6 volts forward bias and generally about 10 milliamperes. The life expectancy of the LED is very long, over 100,000 hours of operation.

Figure 3-27.—LED.

LEDs are used widely as "power on" indicators of current and as displays for pocket calculators, digital voltmeters, frequency counters, etc. For use in calculators and similar devices, LEDs are typically placed together in seven-segment displays, as shown in figure 3-28 (view A and view B). This display uses seven LED segments, or bars (labeled A through G in the figure), which can be lit in different combinations to form any number from "0" through "9." The schematic, view A, shows a common-anode display. All anodes in a display are internally connected. When a negative voltage is applied to the proper cathodes, a number is formed. For example, if negative voltage is applied to all cathodes except that of LED "E," the number "9" is produced, as shown in view A of figure 3-29. If the negative voltage is changed and applied to all cathodes except LED "B," the number "9" changes to "6" as shown in view B.

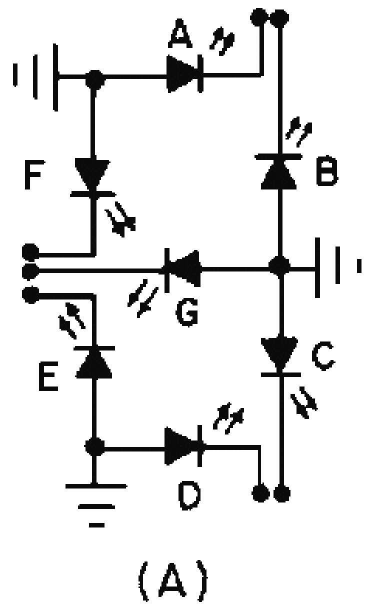


Figure 3-28A.—Seven-segment LED display.

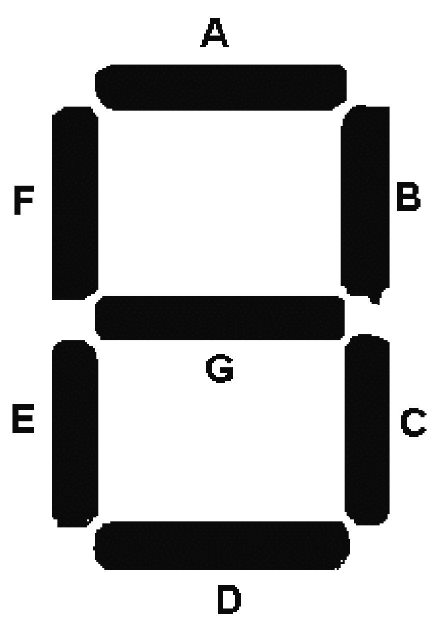


Figure 3-28B.—Seven-segment LED display.

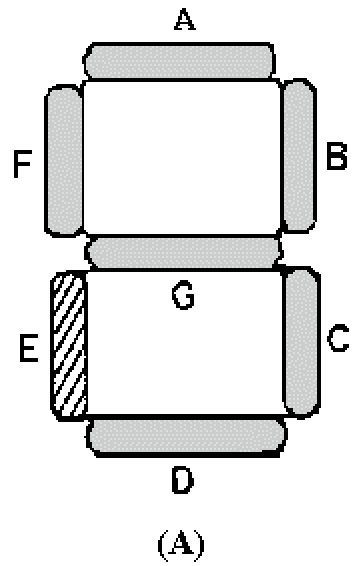


Figure 3-29A.—Seven-segment LED display examples.

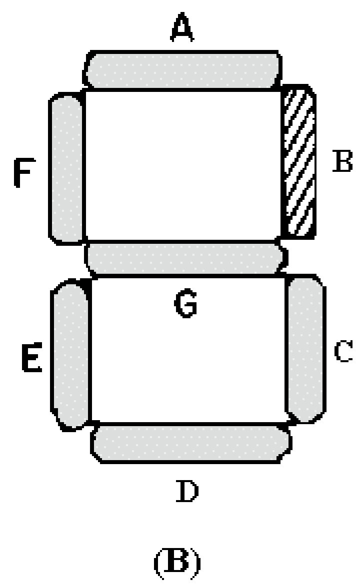


Figure 3-29B.—Seven-segment LED display examples.

Seven-segment displays are also available in common-cathode form, in which all cathodes are at the same potential. When replacing LED displays, you must ensure the replacement display is the same type as the faulty display. Since both types look alike, you should always check the manufacturer's number.

LED seven-segment displays range from the very small, often not much larger than standard typewritten numbers, to about an inch. Several displays may be combined in a package to show a series of numbers, such as the one shown in figure 3-30.

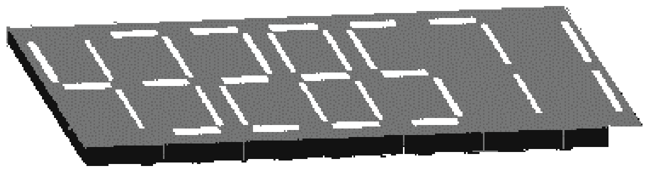


Figure 3-30.—Stacked seven-segment display.

Another special optoelectronic device in common use today is the photodiode. Unlike the LED, which produces light, the photodiode uses light to accomplish special circuit functions. Basically, the PHOTODIODE is a light-controlled variable resistor. In total darkness, it has a relatively high resistance and therefore conducts little current. However, when the PN junction is exposed to an external light source, internal resistance decreases and current flow increases. The photodiode is operated with reverse-bias and conducts current in direct proportion to the intensity of the light source.

Figure 3-31 shows a photodiode with its schematic symbol. The arrows pointing toward the symbol indicate that light is required for operation of the device. A light source is aimed at the photodiode through a transparent "window" placed over the semiconductor chip. Switching the light source on or off changes the conduction level of the photodiode. Varying the light intensity controls the amount of conduction. Because photodiodes respond quickly to changes in light intensity, they are extremely useful in digital applications such as computer card readers, paper tape readers, and photographic light meters. They are also used in some types of optical scanning equipment.

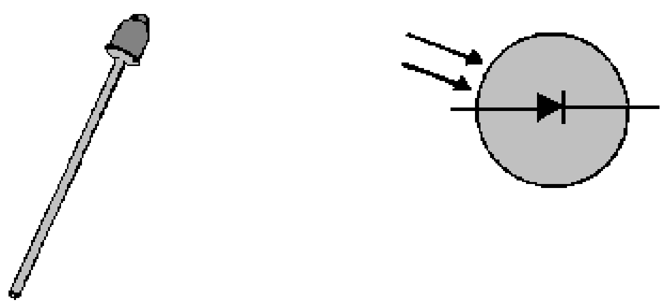


Figure 3-31.—Photodiode.

A second optoelectronic device that conducts current when exposed to light is the PHOTOTRANSISTOR. A phototransistor, however, is much more sensitive to light and produces more output current for a given light intensity that does a photodiode. Figure 3-32 shows one type of phototransistor, which is made by placing a photodiode in the base circuit of an NPN transistor. Light falling on the photodiode changes the base current of the transistor, causing the collector current to be amplified. Phototransistors may also be of the PNP type, with the photodiode placed in the base-collector circuit.

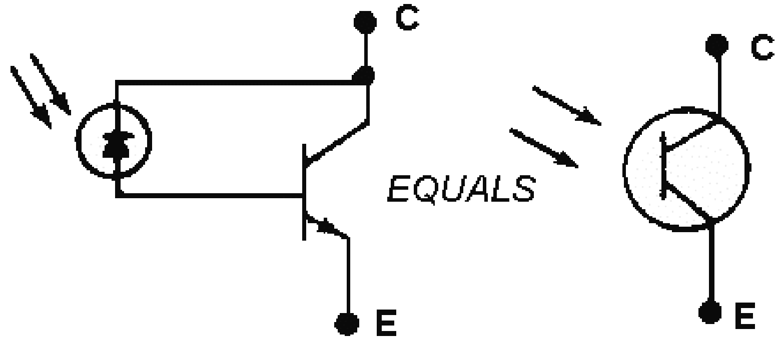


Figure 3-32.—Phototransistor.

Figure 3-33 illustrates the schematic symbols for the various types of phototransistors. Phototransistors may be of the two-terminal type, in which the light intensity on the photodiode alone determines the amount of conduction. They may also be of the three-terminal type, which have an added base lead that allows an electrical bias to be applied to the base. The bias allows an optimum transistor conduction level, and thus compensates for ambient (normal room) light intensity.

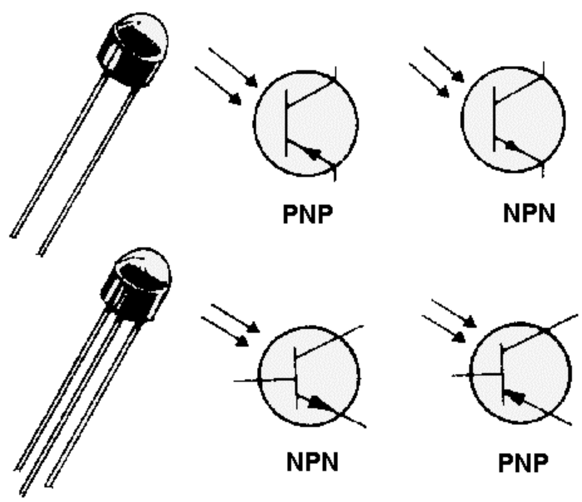


Figure 3-33.—2-terminal and 3-terminal phototransistors.

An older device that uses light in a way similar to the photodiode is the photoconductive cell, or PHOTOCELL, shown with its schematic symbol in figure 3-34. Like the photodiode, the photocell is a light-controlled variable resistor. However, a typical light-to-dark resistance ratio for a photocell is 1:1000. This means that its resistance could range from 1000 ohms in the light to 1000 kilohms in the dark, or from 2000 ohms in the light to 2000 kilohms in the dark, and so forth. Of course, other ratios are also available. Photocells are used in various types of control and timing circuits as, for example, the automatic street light controllers in most cities.

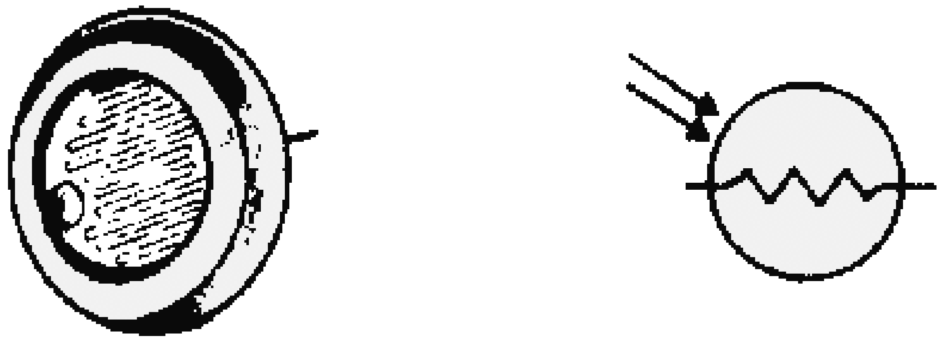


Figure 3-34.—Photocell.

The photovoltaic cell, or solar cell, is a device which converts light energy into electrical energy. An example of a solar cell and its schematic symbol are shown in figure 3-35. The symbol is similar to that of a battery. The device itself acts much like a battery when exposed to light and produces about .45 volt across its terminals, with current capacity determined by its size. As with batteries, solar cells may be connected in series or parallel to produce higher voltages and currents. The device is finding widespread application in communications satellites and solar-powered homes.

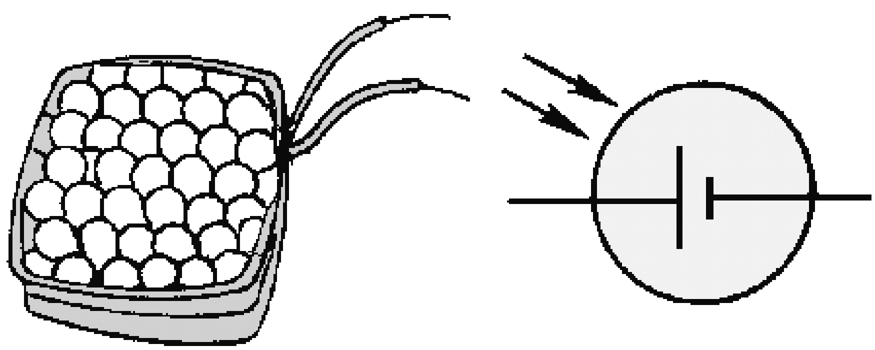


Figure 3-35.—Solar cell.

When it is necessary to block the voltage between one electronic circuit and another, and transfer the signal at the same time, an amplifier coupling capacitor is often used as shown in figure 3-36. Although this method of coupling does block dc between the circuits, voltage isolation is not complete. A newer method, making use of optoelectronic devices to achieve electrical isolation, is the optical coupler, shown in figure 3-37. The coupler is composed of an LED and a photodiode contained in a light-conducting medium. As the polarity signs in figure 3-37 show, the LED is forward biased, while the photodiode is reverse biased. When the input signal causes current through the LED to increase, the light produce by the LED increases. This increased light intensity causes current flow through the photodiode to increase. In this way, changes in input current produce proportional changes in the output, even though the two circuits are electrically isolated.

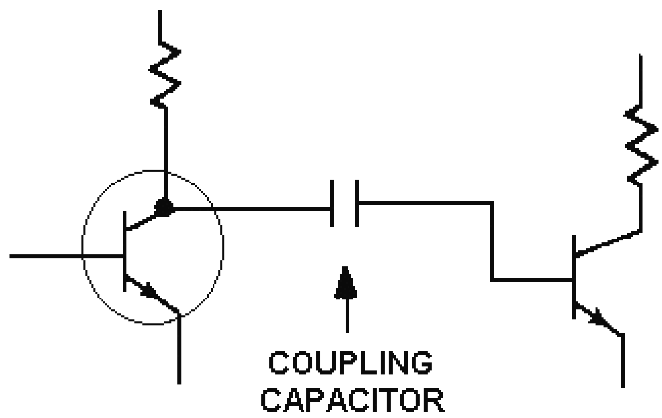


Figure 3-36.—Dc blocking with a coupling capacitor.

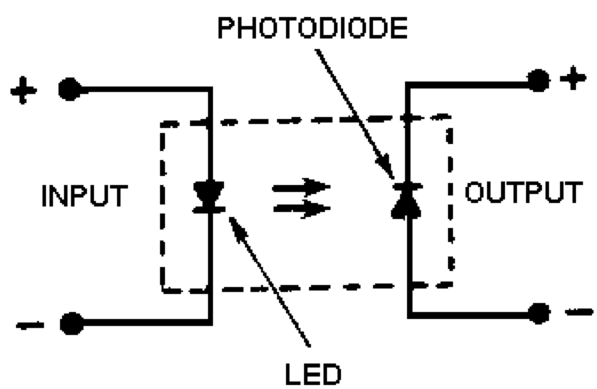


Figure 3-37.—Optical coupler.

The optical coupler is suitable for frequencies in the low megahertz range. The photodiode type shown above can handle only small currents; however, other types of couplers, combining phototransistors with the SCR, can be used where more output is required. Optical couplers are replacing transformers in low-voltage and low-current applications. Sensitive digital circuits can use the coupler to control large current and voltages with low-voltage logic levels.

Q19. What type of bias is required to cause an LED to produce light?

Q20. When compared to incandescent lamps, what is the power requirement of an LED?

Q21. In a common anode, seven-segment LED display, an individual LED will light if a negative voltage is applied to what element?

Q22. What is the resistance level of a photodiode in total darkness?

Q23. What type of bias is required for proper operation of a photodiode?

Q24. What is a typical light-to-dark resistance ratio for a photocell?

Q25. What semiconductor device produces electrical energy when exposed to light?

TRANSISTORS

Transistors are semiconductor devices with three or more terminals. The operation of normal transistors has already been discussed, but there are several transistors with special properties that should be explained. As with diodes, a discussion of all the developments in the transistor field would be impossible. The unijunction transistor (UJT) and the field effect transistor (FET) will be discussed because of their widespread application in Navy equipment. Many other special transistors have been developed and will be discussed in later *NEETS* modules.

The Unijunction Transistor (UJT)

The UNIJUNCTION TRANSISTOR (UJT), originally called a double-based diode, is a three-terminal, solid-state device that has several advantages over conventional transistors. It is very stable over a wide range of temperatures and allows a reduction of components when used in place of conventional transistors. A comparison is shown in figure 3-38. View A is a circuit using conventional transistors, and view B is the same circuit using the UJT. As you can see, the UJT circuit has fewer components. Reducing the number of components reduces the cost, size, and probability of failure.

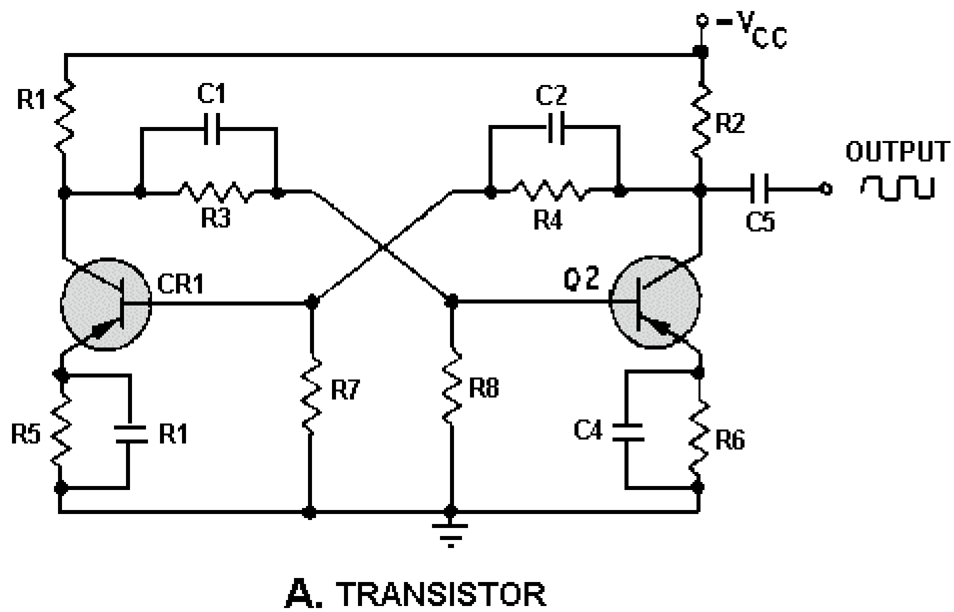


Figure 3-38A.—Comparison of conventional transistors and UJT circuits.

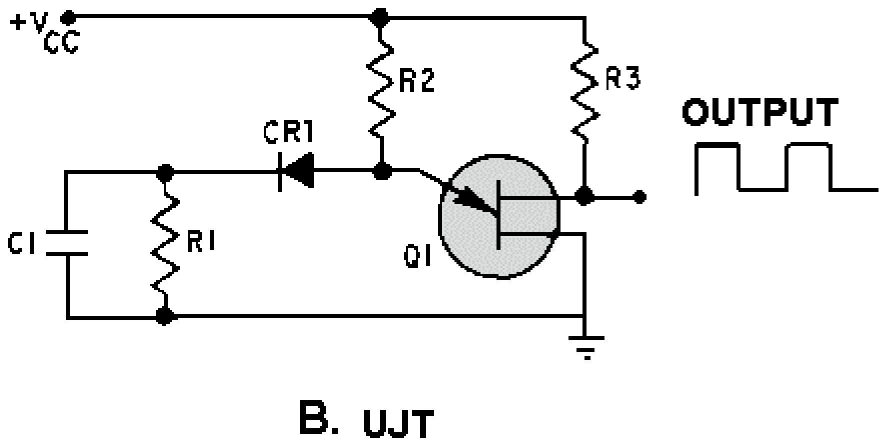


Figure 3-38B.—Comparison of conventional transistors and UJT circuits.

The physical appearance of the UJT is identical to that of the common transistor. As shown in figure 3-39, both have three leads and the same basic shape; the tab on the case indicates the emitter on both devices. The UJT, however, has a second base instead of a collector.

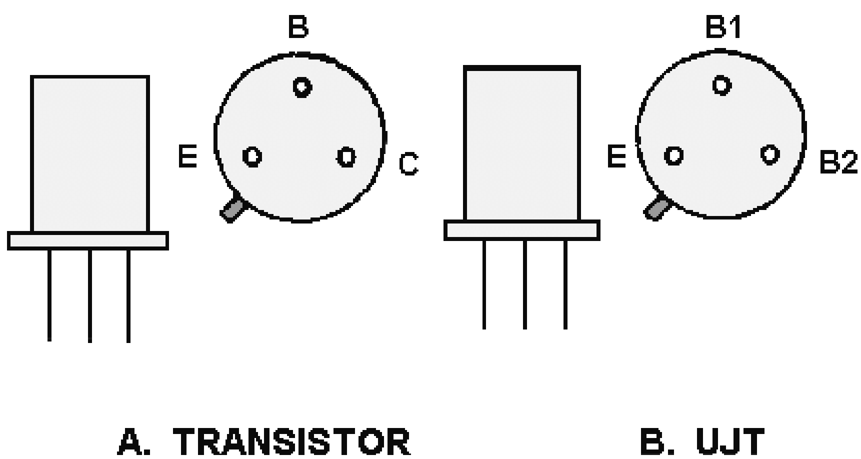


Figure 3-39.—Transistor and UJT.

As indicated in the block diagram shown in views A and B of figure 3-40, the lead differences are even more pronounced. Unlike the transistor, the UJT has only one PN junction. The area between base 1 and base 2 acts as a resistor when the UJT is properly biased. A conventional transistor needs a certain bias level between the emitter, base, and collector for proper conduction. The same principle is true for the UJT; it needs a certain bias level between the emitter and base 1 and also between base 1 and base 2 for proper conduction.

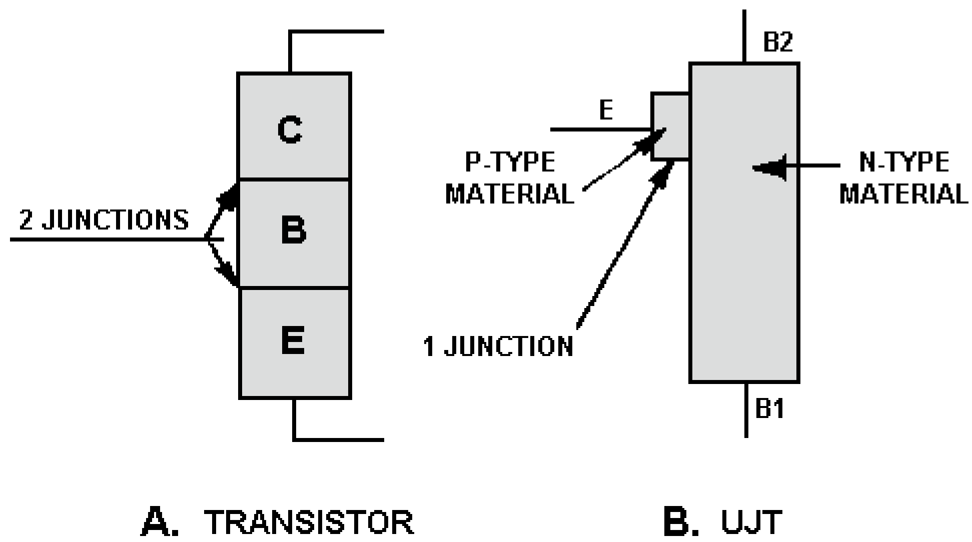


Figure 3-40.—Transistor and UJT structure.

The normal bias arrangement for the UJT is illustrated in figure 3-41, view A. A positive 10 volts is placed on base 2 and a ground on base 1. The area between base 1 and base 2 acts as a resistor. If a reading were taken between base 1 and base 2, the meter would indicate the full 10 volts as shown in view B. Theoretically, if one meter lead were connected to base 1 and the other lead to some point between base 1 and base 2, the meter would read some voltage less than 10 volts. This concept is illustrated in figure 3-42, view A. View B is an illustration of the voltage levels at different points between the two bases. The sequential rise in voltage is called a voltage gradient.

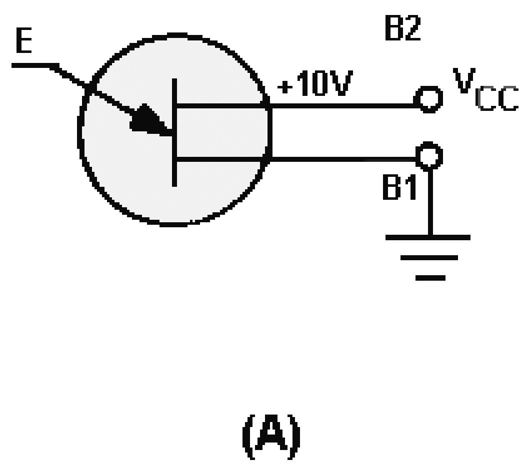


Figure 3-41A.—UJT biasing.

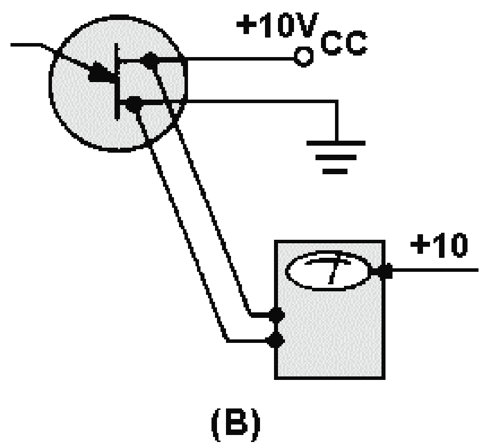


Figure 3-41B.—UJT biasing.

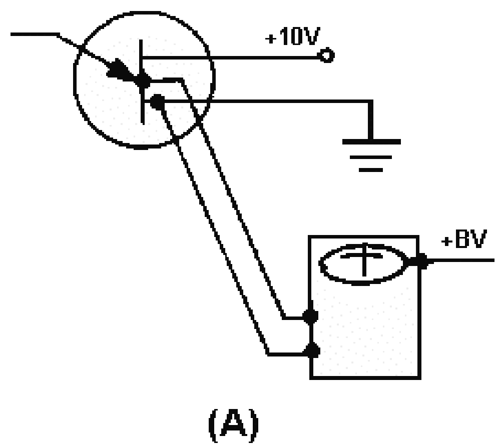


Figure 3-42A.—UJT voltage gradient.

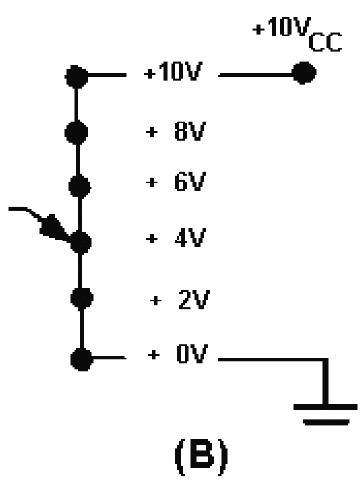


Figure 3-42B.—UJT voltage gradient.

The emitter of the UJT can be viewed as the wiper arm of a variable resistor. If the voltage level on the emitter is more positive than the voltage gradient level at the emitter-base material contact point, the UJT is forward biased. The UJT will conduct heavily (almost a short circuit) from base 1 to the emitter. The emitter is fixed in position by the manufacturer. The level of the voltage gradient therefore depends upon the amount of bias voltage, as shown in figure 3-43.

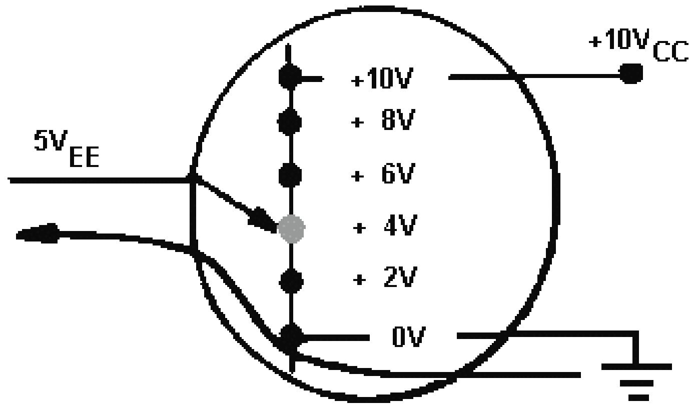


Figure 3-43.—Forward bias point on UJT voltage gradient.

If the voltage level on the emitter is less positive than the voltage gradient opposite the emitter, the UJT is reverse biased. No current will flow from base 1 to the emitter. However, a small current, called reverse current, will flow from the emitter to base 2. The reverse current is caused by the impurities used in the construction of the UJT and is in the form of minority carriers.

More than 40 distinct types of UJTs are presently in use. One of the most common applications is in switching circuits. They are also used extensively in oscillators and wave-shaping circuits.

Q26. The UJT has how many PN junctions?

Q27. The area between base 1 and base 2 in a UJT acts as what type of common circuit component?

Q28. The sequential rise in voltage between the two bases of the UJT is called what?

Q29. What is the normal current path for a UJT?

Field Effect Transistors

Although it has brought about a revolution in the design of electronic equipment, the bipolar (PNP/NPN) transistor still has one very undesirable characteristic. The low input impedance associated with its base-emitter junction causes problems in matching impedances between interstage amplifiers.

For years, scientists searched for a solution that would combine the high input impedance of the vacuum tube with the many other advantages of the transistor. The result of this research is the FIELD-EFFECT TRANSISTOR (FET). In contrast to the bipolar transistor, which uses bias current between base and emitter to control conductivity, the FET uses voltage to control an electrostatic field within the transistor. Because the FET is voltage-controlled, much like a vacuum tube, it is sometimes called the "solid-state vacuum tube."

The elements of one type of FET, the junction type (JFET), are compared with the bipolar transistor and the vacuum tube in figure 3-44. As the figure shows, the JFET is a three-element device comparable to the other two. The "gate" element of the JFET corresponds very closely in operation to the base of the transistor and the grid of the vacuum tube. The "source" and "drain" elements of the JFET correspond to the emitter and collector of the transistor and to the cathode and plate of the vacuum tube.

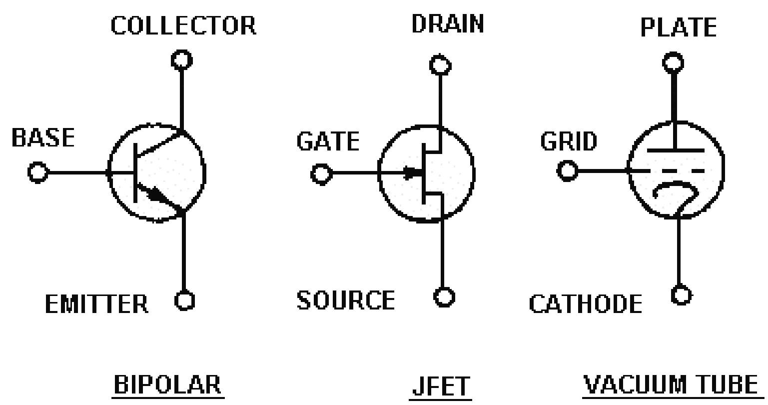


Figure 3-44.—Comparison of JFET, transistor, and vacuum tube symbols.

The construction of a JFET is shown in figure 3-45. A solid bar, made either of N-type or P-type material, forms the main body of the device. Diffused into each side of this bar are two deposits of material of the opposite type from the bar material, which form the "gate." The portion of the bar between the deposits of gate material is of a smaller cross section than the rest of the bar and forms a "channel" connecting the source and the drain. Figure 3-45 shows a bar of N-type material and a gate of P-type material. Because the material in the channel is N-type, the device is called an N-channel JFET.

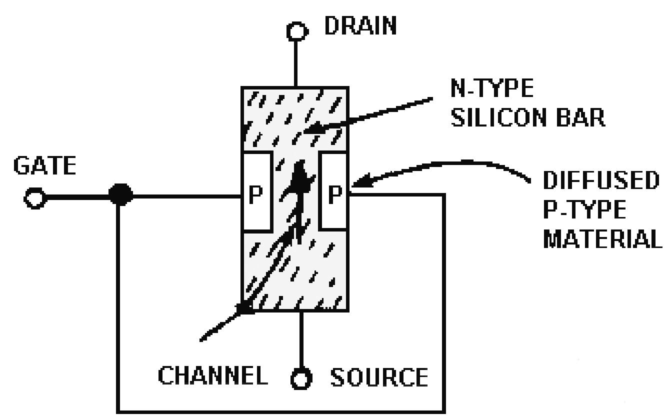


Figure 3-45.—JFET structure.

In a P-channel JFET, the channel is made of P-type material and the gate of N-type material. In figure 3-46, schematic symbols for the two types of JFET are compared with those of the NPN and PNP bipolar transistors. Like the bipolar transistor types, the two types of JFET differ only in the configuration of bias voltages required and in the direction of the arrow within the symbol. Just as it does in transistor symbols, the arrow in a JFET symbol always points towards the N-type material. Thus the symbol of the N-channel JFET shows the arrow pointing toward the drain/source channel, whereas the P-channel symbol shows the arrow pointing away from the drain/source channel toward the gate.

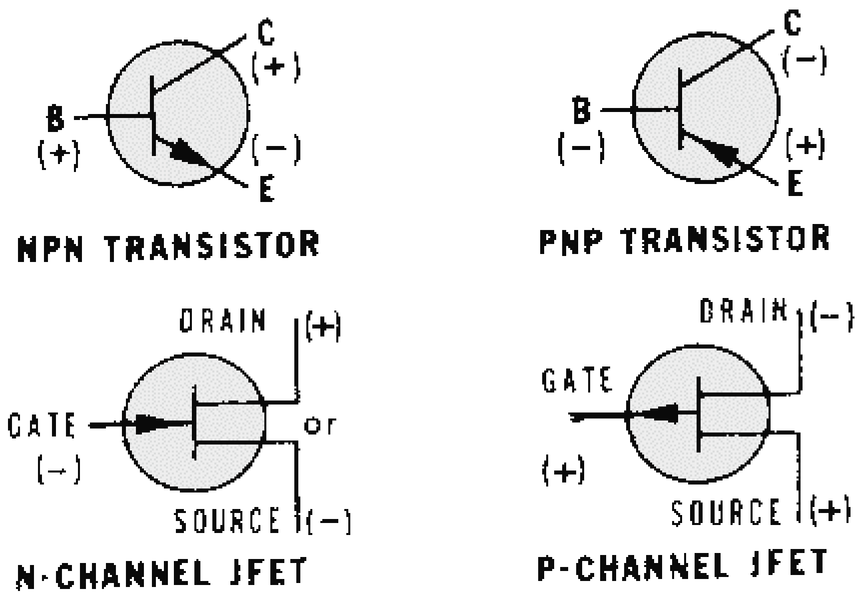


Figure 3-46.—Symbols and bias voltages for transistors and JFET.

The key to FET operation is the effective cross-sectional area of the channel, which can be controlled by variations in the voltage applied to the gate. This is demonstrated in the figures which follow.

Figure 3-47 shows how the JFET operates in a zero gate bias condition. Five volts are applied across the JFET so that current flows through the bar from source to drain, as indicated by the arrow. The gate terminal is tied to ground. This is a zero gate bias condition. In this condition, a typical bar represents a resistance of about 500 ohms. A milliammeter, connected in series with the drain lead and dc power, indicates the amount of current flow. With a drain supply (V_{DD}) of 5 volts, the milliammeter gives a drain current (I_D) reading of 10 milliamperes. The voltage and current subscript letters (V_{DD}, I_D) used for an FET correspond to the elements of the FET just as they do for the elements of transistors.

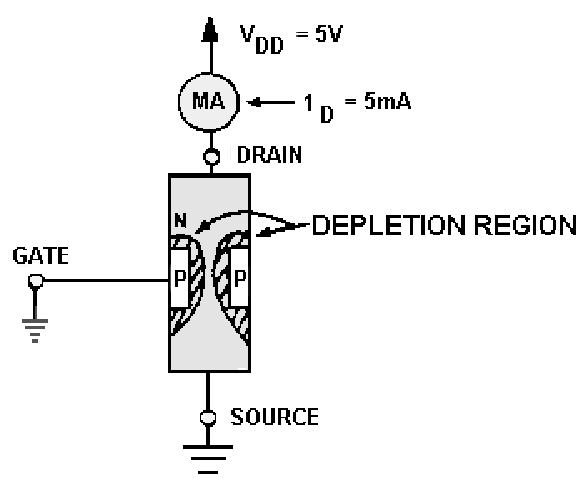


Figure 3-47.—JFET operation with zero gate bias.

In figure 3-48, a small reverse-bias voltage is applied to the gate of the JFET. A gate-source voltage (V_{GG}) of negative 1 volt applied to the P-type gate material causes the junction between the P- and N-type material to become reverse biased. Just as it did in the varactor diode, a reverse-bias condition causes a "depletion region" to form around the PN junction of the JFET. Because this region has a reduced number of current carriers, the effect of reverse biasing is to reduce the effective cross-sectional area of the "channel." This reduction in area increases the source-to-drain resistance of the device and decreases current flow.

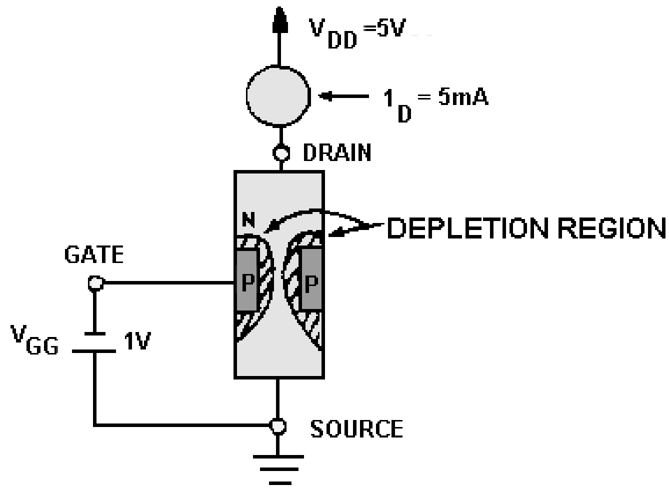


Figure 3-48.—JFET with reverse bias.

The application of a large enough negative voltage to the gate will cause the depletion region to become so large that conduction of current through the bar stops altogether. The voltage required to reduce drain current (I_D) to zero is called "pinch-off" voltage and is comparable to "cut-off" voltage in a vacuum tube. In figure 3-48, the negative 1 volt applied, although not large enough to completely stop conduction, has caused the drain current to decrease markedly (from 10 milliamperes under zero gate bias conditions to 5 milliamperes). Calculation shows that the 1-volt gate bias has also increased the resistance of the JFET (from 500 ohms to 1 kilohm). In other words, a 1-volt change in gate voltage has doubled the resistance of the device and cut current flow in half.

These measurements, however, show only that a JFET operates in a manner similar to a bipolar transistor, even though the two are constructed differently. As stated before, the main advantage of an FET is that its input impedance is significantly higher than that of a bipolar transistor. The higher input impedance of the JFET under reverse gate bias conditions can be seen by connecting a microammeter in series with the gate-source voltage (V_{GG}), as shown in figure 3-49.

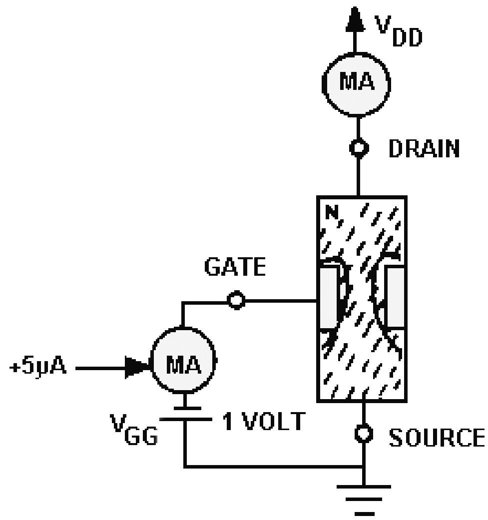


Figure 3-49.—JFET input impedance.

With a V_{GG} of 1 volt, the microammeter reads .5 microamps. Applying Ohm's law (1V ÷ .5 µA) illustrates that this very small amount of current flow results in a very high input impedance (about 2 megohms). By contrast, a bipolar transistor in similar circumstances would require higher current flow (e.g., .1 to −1 mA), resulting in a much lower input impedance (about 1000 ohms or less). The higher input impedance of the JFET is possible because of the way reverse-bias gate voltage affects the cross-sectional area of the channel.

The preceding example of JFET operation uses an N-channel JFET. However, a P-channel JFET operates on identical principles. The differences between the two types are shown in figure 3-50.

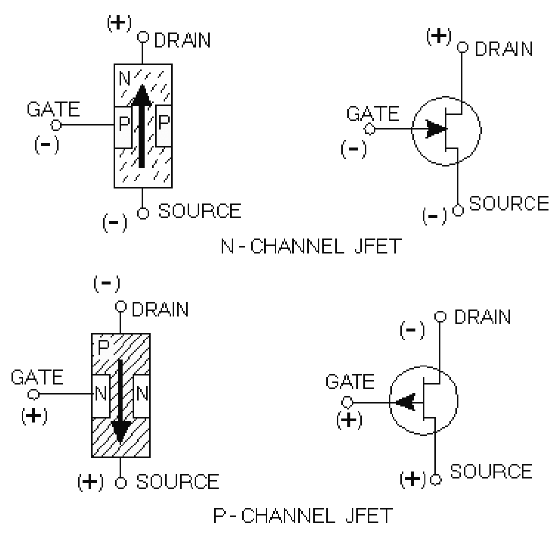


Figure 3-50.—JFET symbols and bias voltages.

Because the materials used to make the bar and the gate are reversed, source voltage potentials must also be reversed. The P-channel JFET therefore requires a positive gate voltage to be reverse biased, and current flows through it from drain to source.

Figure 3-51 shows a basic common-source amplifier circuit containing an N-channel JFET. The characteristics of this circuit include high input impedance and a high voltage gain. The function of the circuit components in this figure is very similar to those in a triode vacuum tube common-cathode amplifier circuit. C1 and C3 are the input and output coupling capacitors. R1 is the gate return resistor and functions much like the grid return resistor in a vacuum tube circuit. It prevents unwanted charge buildup on the gate by providing a discharge path for C1. R2 and C2 provide source self-bias for the JFET, which operates like cathode self-bias. R3 is the drain load resistor, which acts like the plate or collector load resistor.

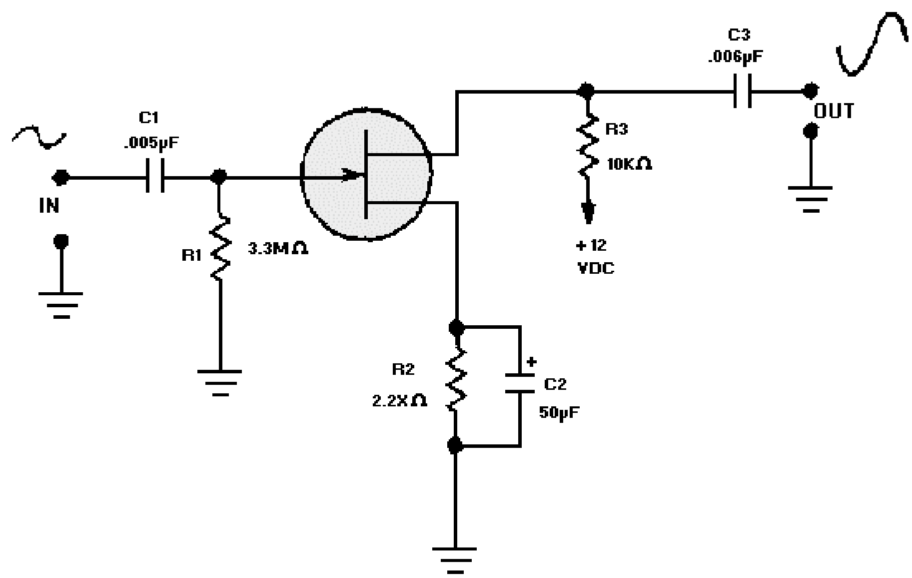


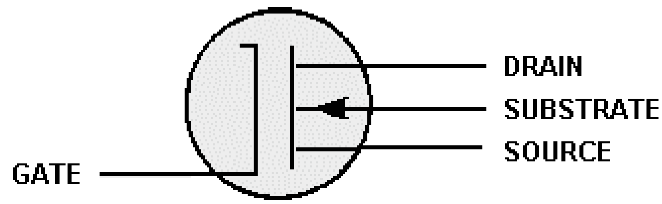
Figure 3-51.—JFET common source amplifier.

The phase shift of 180 degrees between input and output signals is the same as that of common-cathode vacuum tube circuits (and common-emitter transistor circuits). The reason for the phase shift can be seen easily by observing the operation of the N-channel JFET. On the positive alternation of the input signal, the amount of reverse bias on the P-type gate material is reduced, thus increasing the effective cross-sectional area of the channel and decreasing source-to-drain resistance. When resistance decreases, current flow through the JFET increases. This increase causes the voltage drop across R3 to increase, which in turn causes the drain voltage to decrease. On the negative alternation of the cycle, the amount of reverse bias on the gate of the JFET is increased and the action of the circuit is reversed. The result is an output signal, which is an amplified 180-degree-out-of-phase version of the input signal.

A second type of field-effect transistor has been introduced in recent years that has some advantages over the JFET. This device is the metal oxide semiconductor field effect transistor (MOSFET). The MOSFET has an even higher input impedance than the JFET (10 to 100 million megohms). Therefore, the MOSFET is even less of a load on preceding circuits. The extremely high input impedance, combined with a high gain factor, makes the MOSFET a highly efficient input device for RF/IF amplifiers and mixers and for many types of test equipment.

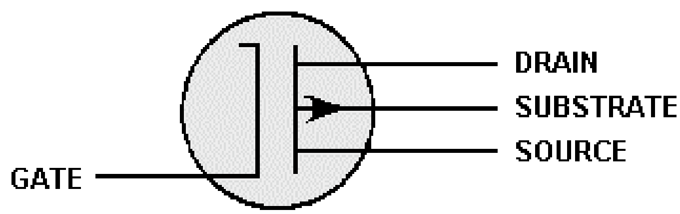
The MOSFET is normally constructed so that it operates in one of two basic modes: the depletion mode or the enhancement mode. The depletion mode MOSFET has a heavily doped channel and uses reverse bias on the gate to cause a depletion of current carriers in the channel. The JFET also operates in this manner. The enhancement mode MOSFET has a lightly doped channel and uses forward bias to enhance the current carriers in the channel. A MOSFET can be constructed that will operate in either mode depending upon what type of bias is applied, thus allowing a greater range of input signals.

In addition to the two basic modes of operation, the MOSFET, like the JFET, is of either the P-channel type or the N-channel type. Each type has four elements: gate, source, drain, and substrate. The schematic symbols for the four basic variations of the MOSFET are shown in views A, B, C, and D of figure 3-52.



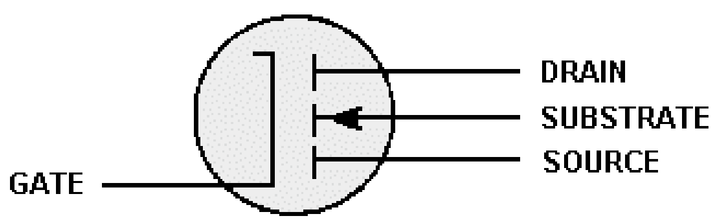
A. N-CHANNEL, DEPLETION, MOSFET

Figure 3-52A.—MOSFET symbols.



B. P-CHANNEL, DEPLETION, MOSFET

Figure 3-52B.—MOSFET symbols.



C. N-CHANNEL, ENHANCEMENT, MOSFET

Figure 3-52C.—MOSFET symbols.

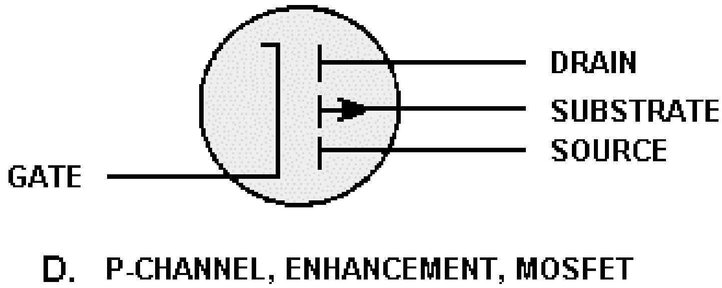


Figure 3-52D.—MOSFET symbols.

The construction of an N-channel MOSFET is shown in figure 3-53. Heavily doped N-type regions (indicated by the N+) are diffused into a P-type substrate or base. A channel of regular N-type material is diffused between the heavily doped N-type regions. A metal oxide insulating layer is then formed over the channel, and a metal gate layer is deposited over the insulating layer. There is no electrical connection between the gate and the rest of the device. This construction method results in the extremely high input impedance of the MOSFET. Another common name for the device, derived from the construction method, is the insulated gate field effect transistor (IGFET).

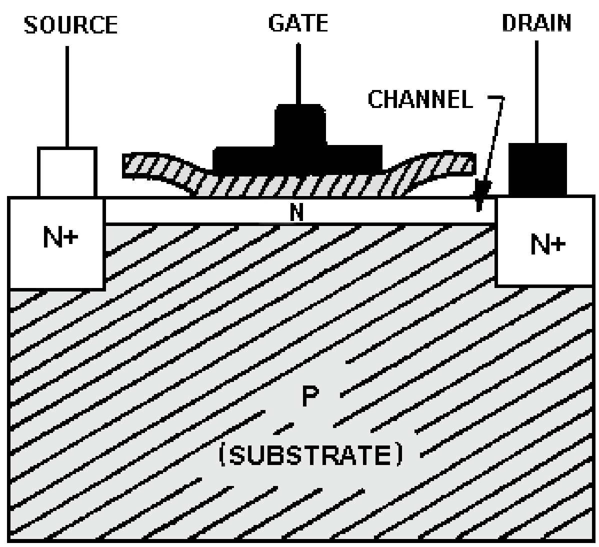


Figure 3-53.—MOSFET structure.

The operation of the MOSFET, or IGFET, is basically the same as the operation of the JFET. The current flow between the source and drain can be controlled by using either of two methods or by using a combination of the two methods. In one method the drain voltage controls the current when the gate potential is at zero volts. A voltage is applied to the gate in the second method. An electric field is formed by the gate voltage that affects the current flow in the channel by either depleting or enhancing the number of current carriers available. As previously stated, a reverse bias applied to the gate depletes the carriers, and a forward bias enhances the carriers. The polarity of the voltages required to forward or reverse bias a MOSFET depends upon whether it is of the P-channel type or the N-channel type. The effects of reverse-bias voltage on a MOSFET designed to operate in the depletion mode are illustrated in views A, B, and C of figure 3-54. The amount of reverse bias applied has a direct effect on the width of the current channel and, thus, the amount of drain current (I_D).

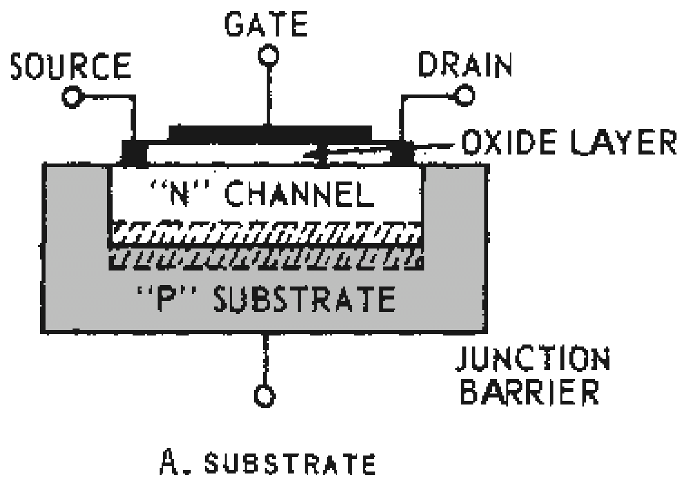


Figure 3-54A.—Effects of bias on N-channel depletion MOSFET.

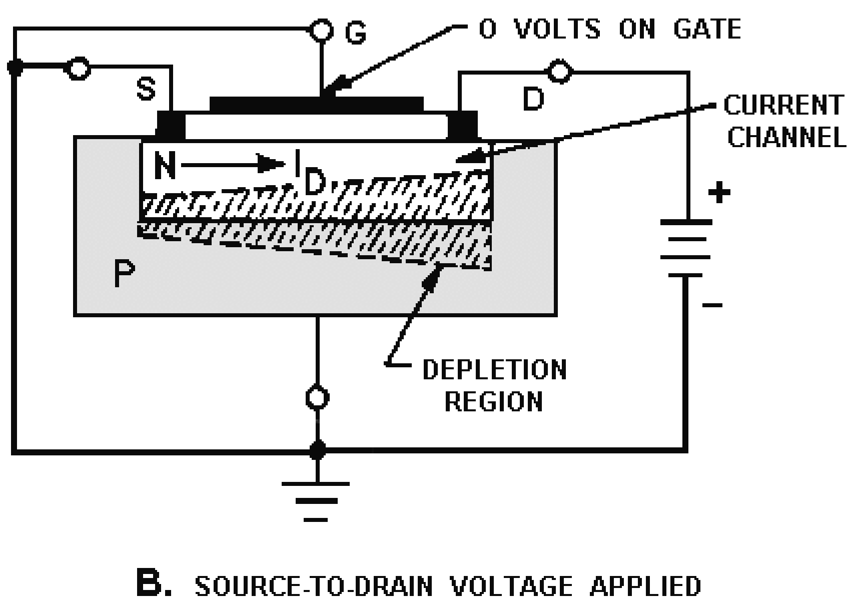
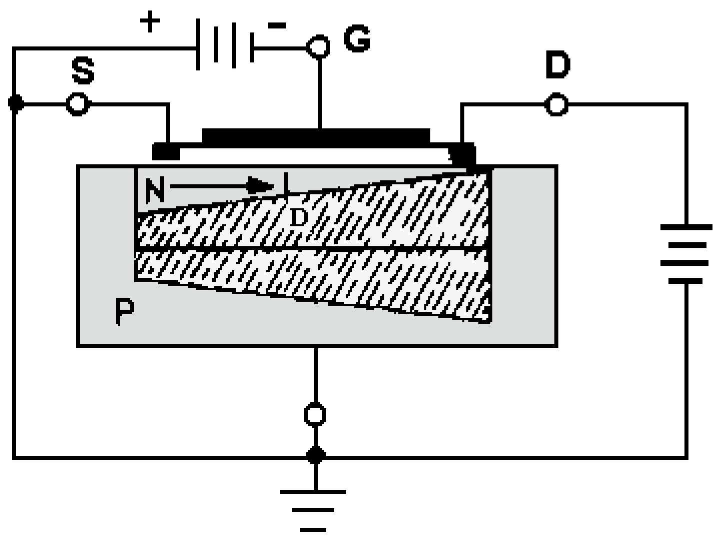


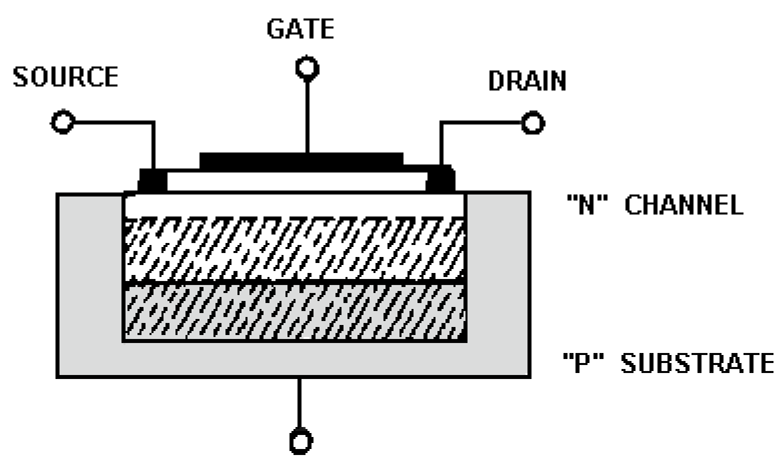
Figure 3-54B.—Effects of bias on N-channel depletion MOSFET.



C. REVERSE BIAS APPLIED

Figure 3-54C.—Effects of bias on N-channel depletion MOSFET.

Figure 3-55 (view A, view B, and view C) illustrates the effect of forward bias on an enhancement mode N-channel MOSFET. In this case, a positive voltage applied to the gate increases the width of the current channel and the amount of drain current (I_D).



A. SUBSTRATE

Figure 3-55A.—Effects of bias on N-channel enhancement MOSFET.

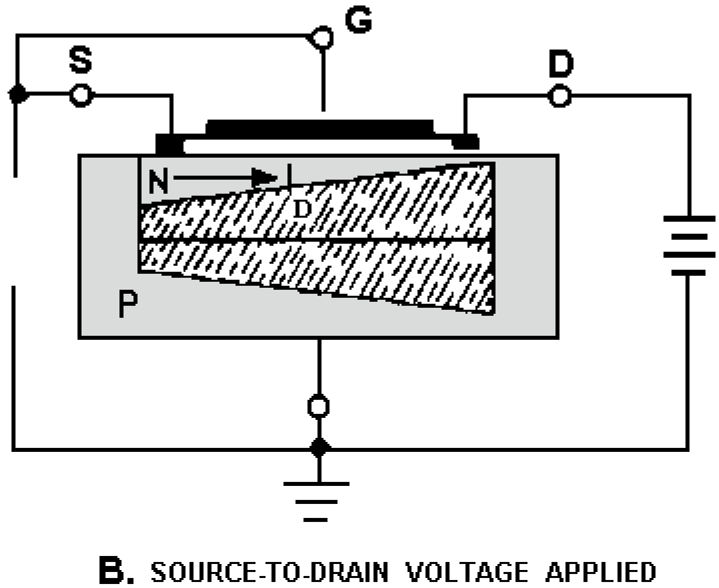


Figure 3-55B.—Effects of bias on N-channel enhancement MOSFET.

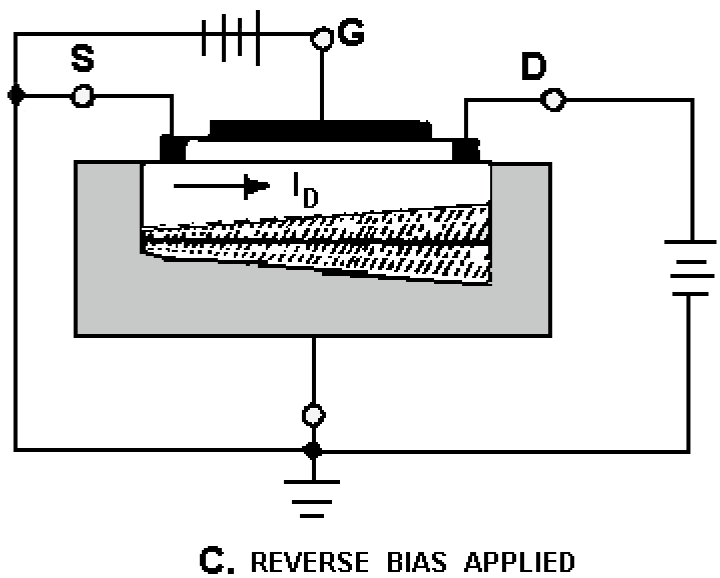


Figure 3-55C.—Effects of bias on N-channel enhancement MOSFET.

Another type of MOSFET is the induced-channel type MOSFET. Unlike the MOSFETs discussed so far, the induced-channel type has no actual channel between the source and the drain. The induced channel MOSFET is constructed by making the channel of the same type material as the substrate, or the opposite of the source and the drain material. As shown in figure 3-56, the source and the drain are of P-type material, and the channel and the substrate are of N-type material.

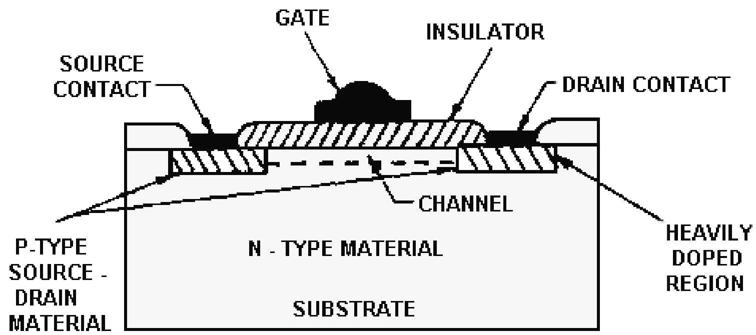


Figure 3-56.—Induced channel MOSFET construction.

The induced-channel MOSFET is caused to conduct from source to drain by the electric field that is created when a voltage is applied to the gate. For example, assume that a negative voltage is applied to the MOSFET in figure 3-56. The effect of the negative voltage modifies the conditions in the substrate material. As the gate builds a negative charge, free electrons are repelled, forming a depletion region. Once a certain level of depletion has occurred (determined by the composition of the substrate material), any additional gate bias attracts positive holes to the surface of the substrate. When enough holes have accumulated at the surface channel area, the channel changes from an N-type material to a P-type material, since it now has more positive carriers than negative carriers. At this point the channel is considered to be to inverted, and the two P-type regions at the source and the drain are now connected by a P-type inversion layer or channel. As with the MOSFET, the gate signal determines the amount of current flow through the channel as long as the source and drain voltages remain constant. When the gate voltage is at zero, essentially no current flows since a gate voltage is required to form a channel.

The MOSFETs discussed up to this point have been single-gate MOSFETs. Another type of MOSFET, the dual-gate type, is shown in figure 3-57. As the figure shows, the gates in a dual-gate MOSFET can be compared to the grids in a multi-grid vacuum tube. Because the substrate has been connected directly to the source terminal, the dual-gate MOSFET still has only four leads: one each for source and drain, and two for the gates. Either gate can control conduction independently, making this type of MOSFET a truly versatile device.

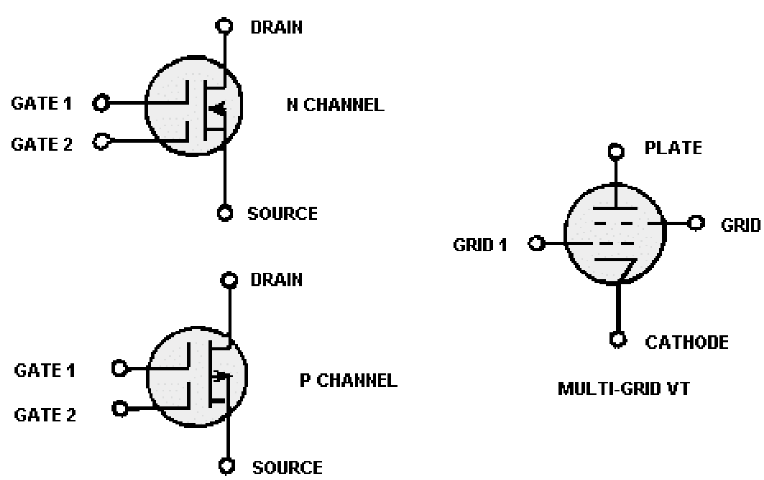


Figure 3-57.—Dual-gate MOSFET.

One problem with both the single- and dual-gate MOSFETs is that the oxide layer between gate and channel can be destroyed very easily by ordinary static electricity. Replacement MOSFETs come packaged with their leads shorted together by a special wire loop or spring to avoid accidental damage. The rule to remember with these shorting springs is that they must not be removed until after the MOSFET has been soldered or plugged into a circuit. One such spring is shown in figure 3-58.

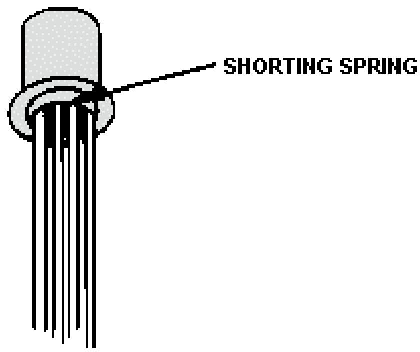


Figure 3-58.—MOSFET shorting spring.

Q30. What is one of the primary advantages of the FET when compared to the bipolar transistor?

Q31. The FET and the vacuum tube have what in common?

Q32. The base of a transistor serves a purpose similar to what element of the FET?

Q33. What are the two types of JFET?

Q34. The source and drain of an N-channel JFET are made of what type of material?

Q35. What is the key to FET operation?

Q36. What is the normal current path in an N-channel JFET?

Q37. Applying a reverse bias to the gate of an FET has what effect?

Q38. The input and output signals of a JFET amplifier have what phase relationship?

Q39. When compared to the JFET, what is the input impedance of the MOSFET?

Q40. What are the four elements of the MOSFET?

Q41. The substrate of an N-channel MOSFET is made of what material?

Q42. In a MOSFET, which element is insulated from the channel material?

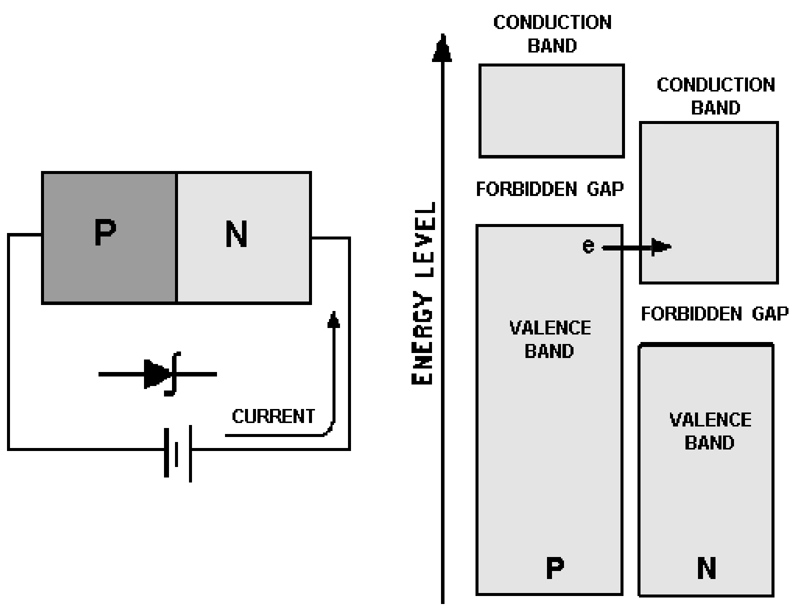
Q43. What type of MOSFET can be independently controlled by two separate signals?

Q44. What is the purpose of the spring or wire around the leads of a new MOSFET?

SUMMARY

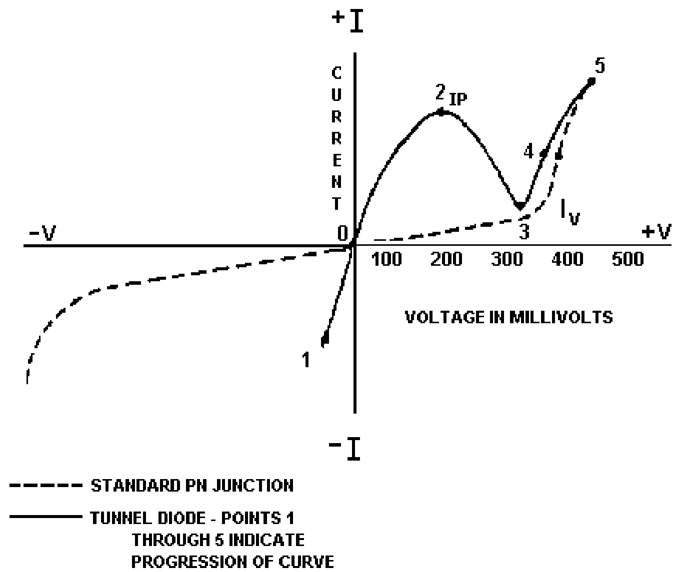
This chapter introduced you to a representative selection of solid-state devices that have special properties. The basic operating principles of the devices discussed in this chapter are summarized in the following paragraphs for you to use as a review and a future reference.

The **ZENER DIODE** is a PN junction that is designed to operate in the reverse-bias breakdown mode. When the applied voltage reaches the breakdown point, the Zener diode, for all practical purposes, becomes a short circuit. The reverse bias and breakdown mode of operation cause the Zener diode to conduct with (in the direction of) the arrow in the symbol as shown.

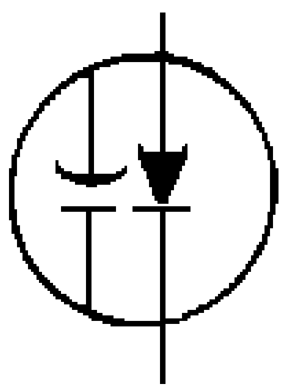


Two theories are used to explain the breakdown action of Zener diodes. The **ZENER EFFECT** explains the breakdown of diodes below 5 volts. The heavy doping used in these diodes allows the valence band of one material to overlap the energy level of the conduction band of the other material. This situation allows electrons to tunnel across the PN junction at the point where the two energy bands overlap. Zener diodes that operate above 5 volts are explained by the AVALANCHE EFFECT in which free electrons colliding with bound electrons cause an ever-increasing number of free current carriers in a multiplying action. The Zener diode is used primarily as a voltage regulator in electronic circuits.

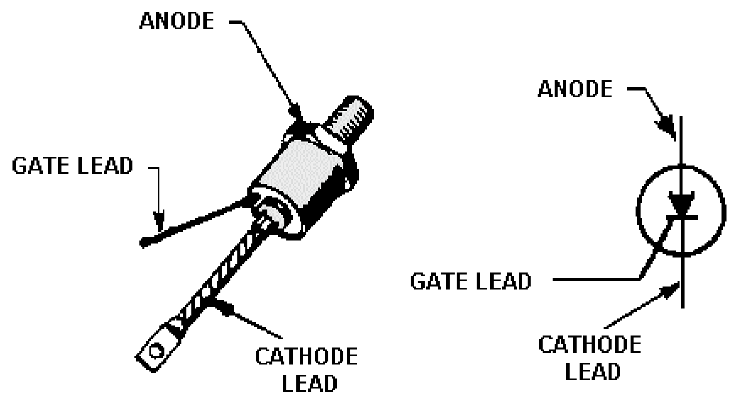
The **TUNNEL DIODE** is a heavily doped PN junction that exhibits negative resistance over part of its range of operation, as can be seen in the curve in the illustration. The heavy doping causes the tunnel diode to have a very narrow depletion region and also causes the valence band of one of the semiconductor materials to overlap the energy level of the conduction band of the other semiconductor material. At the energy overlap point, electrons can cross from the valence band of one material to the conduction band of the other material without acquiring any additional energy. This action is called tunneling. Tunnel diodes are used as amplifiers, oscillators, and high-speed switching devices.



The **VARACTOR** is a diode that exhibits the characteristics of a variable capacitor. The depletion region at the PN junction acts as the dielectric of a capacitor and is caused to expand and contract by the voltage applied to the diode. This action increases and decreases the capacitance. The schematic symbol for the varactor is shown below. Varactors are used in tuning circuits and can be used as high-frequency amplifiers.

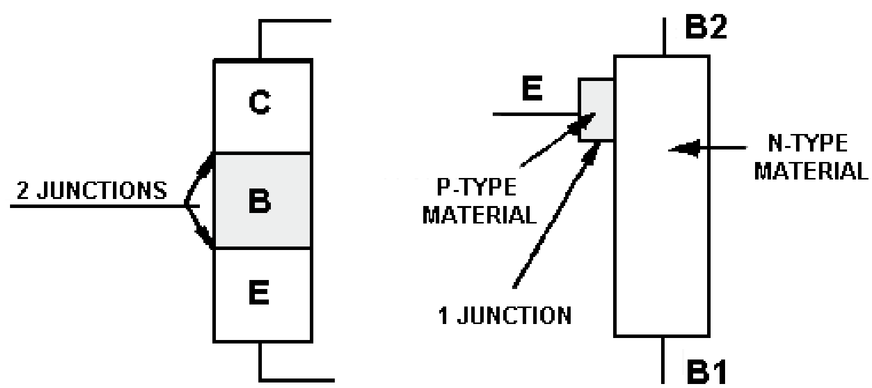


The **SILICON CONTROLLED RECTIFIER (SCR)** is a four-element, solid-state device that combines characteristics of both diodes and transistors. The symbol for the SCR is shown below. A signal must be applied to the gate to cause the SCR to conduct. When the proper gate signal is applied, the SCR conducts or "fires" until the bias potential across the device drops below the minimum required to sustain current flow. Removal of the gate signal does not shut off the SCR. In fact, the gate signal is often a very narrow voltage pulse or trigger. The SCR is ideal for use in situations where a small, low-power gate can be used to turn on larger currents, such as those found in rectifier and switching circuits. SCRs are used extensively in power supply circuits as rectifiers.

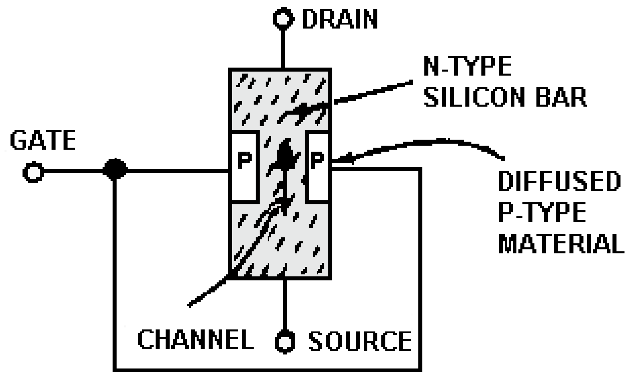


OPTOELECTRONIC DEVICES are of two basic types: light producers or light users. The LED is the most widely used light-producing device. When the LED is forward biased it emits energy in the form of light. LEDs are used in several configurations as digital equipment readout displays. The PHOTODIODE, the PHOTOTRANSISTOR, and the PHOTOCELL are all devices that use light to modify conduction through them. The SOLAR CELL uses light to produce voltage.

The **UNIJUNCTION TRANSISTOR (UJT)** is a three-terminal, solid-state device with only one PN junction. The block diagram below shows the difference in construction between normal transistors and the UJT. The area between base 1 and base 2 of the UJT acts as a variable resister. The emitter of the UJT acts as the wiper arm. The sequential rise in voltage between the bases is called a voltage gradient. The UJT conducts when the emitter is more positive than the voltage gradient at the emitter/base contact point. There are many variations of the UJT which are used in switching circuits, oscillators, and wave-shaping circuits.



The FIELD-EFFECT TRANSISTOR combines the high input impedance of the vacuum tube with all the other advantages of the transistor. The elements of the FET are the gate, source, and drain, which are comparable to the base, emitter, and collector of a standard transistor. The JFET or "junction FET" is made of a solid bar of either P- or N-semiconductor material, and the gate is made of the opposite type material, as illustrated below. The FET is called P-channel or N-channel depending upon the type of material used to make the bar between the source and drain. Voltage applied to the gate controls the width of the channel and consequently controls the current flow from the source to the drain. The JFET is normally operated with reverse bias that controls the channel width by increasing or decreasing the depletion region.



The **MOSFET** is an FET that has even higher input impedances than the JFET because the gate of the MOSFET is completely insulated from the rest of the device. The MOSFET operates in either the depletion mode or the forward-bias enhancement mode and can be either N-channel or P-channel. The induced-channel and the dual-gate MOSFETs are variations of the basic MOSFET.

ANSWERS TO QUESTIONS Q1. THROUGH Q44.

A1. The minority carriers.

A2. Zener effect and avalanche effect.

A3. Zener effect.

A4. The doping level of an avalanche effect diode is lower.

A5. An external current-limiting resistor.

A6. Because Zener diodes are operated in the reverse bias mode.

A7. The amount of doping.

A8. Negative resistance.

A9. The tunnel diode has a very narrow depletion region.

A10. Minimum.

A11. Variable capacitance.

A12. The depletion region decreases.

A13. Capacitance decreases.

A14. The SCR is primarily used for switching power on or off.

A15. A gate signal.

A16. The forward bias must be reduced below the minimum conduction level.

A17. SCR.

A18. During both alternations.

A19. Forward bias.

A20. Very low.

A21. The cathode.

A22. Very high.

A23. Reverse bias.

A24. 1:1000.

A25. Photovoltaic cell.

A26. One.

A27. Variable resistor.

A28. A voltage gradient.

A29. From base 1 to the emitter.

A30. High input impedance.

A31. Voltage controls conduction.

A32. Gate.

A33. N-channel and P-channel.

A34. N-type material.

A35. Effective cross-sectional area of the channel.

A36. From source to drain.

A37. Source-to-drain resistance increases.

A38. They are 180 degrees out of phase.

A39. The MOSFET has a higher input impedance.

A40. Gate, source, drain, and substrate.

A41. P-type material.

A42. The gate terminal.

A43. The dual-gate MOSFET.

A44. To prevent damage from static electricity.